

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

ESTUDIO DE LA VOLATILIDAD ESTOCÁSTICA CON PARÁMETROS VARIABLES EN EL TIEMPO: UNA APLICACIÓN AL PRECIO DEL CRUDO ORIENTE DEL ECUADOR

TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE
INGENIERO MATEMÁTICO

PROYECTO DE INVESTIGACIÓN

PABLO DAVID ALBÁN CERDA

pablodavidac@gmail.com

DIRECTOR: Dr. HOLGER ANÍBAL CAPA SANTOS
holger.capa@epn.edu.ec

QUITO, Julio 2018

DECLARACIÓN

Yo, Pablo David Albán Cerda, declaro bajo juramento que el trabajo aquí descrito es de mí autoría, que no ha sido previamente presentado para ningún grado o calificación personal; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

La escuela Politécnica Nacional puede hacer uso de los derechos correspondientes a este trabajo, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normativa institucional vigente.



Pablo David Albán Cerda

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por Pablo David Albán Cerda, bajo mi supervisión.



Dr. Holger Aníbal Capa Santos

AGRADECIMIENTOS

Gracias a mi familia, quienes me han apoyado durante toda mi vida en especial en mi carrera universitaria.

Gracias al Dr Holger Capa, quien con su acertada dirección y apoyo, me ha permitido presentar este proyecto.

Gracias a la Facultad de Ciencias, por inculcarme todos los conocimientos que me servirán en mi vida profesional.

DEDICATORIA

A mis padres Rubén y Martha, por su amor, sacrificio y ejemplo que me han brindado a lo largo de mi vida académica. Les dedico este trabajo de titulación, en la cual está plasmado no simplemente mi esfuerzo sino también el de ustedes.

A mi hermano Santiago, por su ejemplo y apoyo que me ha brindado durante toda mi vida.

Índice general

1	Introducción	1
1.1	Motivación	1
1.2	Modelo de espacios de estados generalizado	2
1.2.1	El modelo lineal gaussiano de espacio de estados	3
1.3	Modelando la volatilidad de los precios del crudo Oriente	4
1.3.1	Modelos GARCH	4
1.3.2	Modelos con volatilidad estocástica	5
1.3.3	Modelo SVM-TVP	6
1.3.3.1	Modelo SVM-TVP en estado de espacios	7
1.4	Inferencia bayesiana	8
1.4.1	Estimador bayesiano	9
I	Marco Teórico	11
2	Historia e impacto del petróleo en la economía del Ecuador y su incertidumbre	12
2.1	Descubrimiento del petróleo en Ecuador	12
2.2	Relación histórica entre sector petrolero y la economía ecuatoriana	13
2.3	Producción Petrolera del Ecuador	15
2.4	Precios históricos de los crudos Oriente, Napo y WTI	17
2.5	Impacto del precio del Petróleo	18
2.5.1	Impacto en la balanza comercial	18
2.5.2	Impacto en el sector fiscal	20
2.5.2.1	Impacto en el sector público no financiero	20
2.5.2.2	Impacto en la deuda externa	22
2.5.2.3	Impacto en la deuda interna	22
2.6	Incertidumbre en los precios del petróleo	23
3	Métodos de Monte Carlo basados en cadenas de Márkov (MCMC)	26
3.1	Algoritmo de Metropolis-Hasting	27
3.2	Muestreo de Gibbs	28
3.3	Muestreo por aceptación y rechazo (AR)	29
3.4	Algoritmo ARMH	30
4	Estimación bayesiana del modelo SVM-TVP	32
4.1	Muestreo de h	33
4.2	Muestreo de γ	36
4.3	Muestreo de Ω	37

4.4	Muestreo de σ^2	38
4.5	Muestreo de β	38
4.6	Muestreo de μ y ϕ	39
4.6.1	Restricción del parámetro ϕ	40
4.6.2	Cálculo de los parámetros de la función propuesta generadora . . .	40
4.7	Verosimilitud marginal	42
4.7.1	Verosimilitud marginal del modelo SVM-TVP	43
4.7.1.1	Elección de ψ^*	44
4.7.1.2	Estimación de la verosimilitud $p(y \psi^*)$	44
4.7.1.3	Estimación de $p(\psi^* y)$	44
4.8	Probabilidades que $\alpha_t \neq 0$	47
4.9	Criterios de comparación entre modelos	47
4.9.1	Factor bayesiano, BIC y AIC	47
4.10	Pruebas de diagnóstico	48
II	Marco Práctico	50
5	Resultados computacionales	51
5.1	Resultados del modelo SVM-TVP con pronósticos a 12 semanas	53
5.1.1	Hiperparámetros de las distribuciones a priori	53
5.1.2	Estimación del modelo SVM-TVP para el crudo Oriente	55
5.1.3	Probabilidades que $\alpha_t \neq 0$	57
5.1.4	Resultados de las pruebas de diagnóstico	58
5.1.5	Predicciones del modelo SVM-TVP	58
5.1.6	Modelo SVM-TVP con $\alpha_t = 0$	60
5.1.7	Comparación con otros modelos	64
5.2	Resultados del modelo SVM-TVP con pronósticos a 24 semanas	66
5.2.1	Estimación del modelo SVM-TVP para el crudo Oriente	66
5.2.2	Resultados de las pruebas de diagnóstico	67
5.2.3	Predicciones del modelo SVM-TVP	68
5.2.4	Modelo SVM-TVP con $\alpha_t = 0$	70
5.2.5	Comparación con otros modelos	72
6	Conclusiones y recomendaciones	76
A	Código en Matlab	80

Índice de figuras

2.1	Evolución de los precios del crudo ecuatoriano.	14
2.2	Evolución mensual de la producción petrolera ecuatoriana.	15
2.3	Producción petrolera ecuatoriana por empresas.	16
2.4	Producción diaria promedio de Petroamazonas EP y Operadora Río Napo .	16
2.5	Producción diaria promedio de las compañías privadas en miles de barriles.	17
2.6	Evolución de los precios promedio anuales del crudo WTI, Oriente y Napo .	17
2.7	Evolución de los precios promedio mensuales del crudo WTI, Oriente y Napo	18
2.8	Evolución de la balanza comercial versus los precios del petróleo	19
2.9	Evolución de la balanza comercial mensual millones de dólares.	19
2.10	Evolución anual de las exportaciones en millones de dólares.	20
2.11	Evolución del sector público no financiero entre 2014 y 2016.	21
2.12	Ingresos del SPNF entre 2014 y 2016.	21
2.13	Evolución de la deuda externa (pública y privada) en millones de dólares. .	22
2.14	Evolución de la deuda externa e interna del Ecuador.	23
2.15	Evolución del índice de las noticias relacionadas con la OPEP.	24
2.16	Serie VIX versus OVX	24
5.1	Evolución de las series z_t y y_t desde el 2012 al 2017	52
5.2	Primeras 5.000 simulaciones de las f.d.p. a posteriori y verosimilitud	54
5.3	Simulaciones de las f.d.p. a posteriori y verosimilitud	56
5.4	Evolución de las f.d.p. a posteriori de h_t , α_t y τ_t	57
5.5	Probabilidades a posteriori que $\alpha_t \neq 0$	57
5.6	Pruebas de Ljung-Box y Li-McLeod	58
5.7	Predicciones e intervalos de credibilidad, al nivel del 95 %, para los precios	59
5.8	Pruebas de Ljung-Box y Li-McLeod con $\alpha_t = 0$	61
5.9	Simulaciones de las f.d.p. a posteriori y verosimilitud con $\alpha_t = 0$	62
5.10	Predicciones e intervalos de credibilidad, al nivel del 95 %, del modelo con $\alpha_t = 0$	63
5.11	Probabilidades a posteriori que $\alpha_t \neq 0$	67
5.12	Pruebas de Ljung-Box y Li-McLeod	68
5.13	Predicciones e intervalos de credibilidad, al nivel del 95 %, para los precios	69
5.14	Pruebas de Ljung-Box y Li-McLeod con $\alpha_t = 0$	71
5.15	Predicciones e intervalos de credibilidad, al nivel del 95 %, del modelo con $\alpha_t = 0$	71
5.16	Evolución de las desviaciones estándar de los modelos	75

Índice de cuadros

1.1	Lista de modelos GARCH y SV	6
4.1	Tasa de aceptación de μ y ϕ	42
5.1	Tiempos de ejecución de los códigos para los diferentes modelos en Matlab.	52
5.2	Estimación a posteriori de los momentos y cuantiles de los parámetros del modelo SVM-TVP guardando 12 observaciones.	55
5.3	Predicciones e intervalos de credibilidad, al nivel del 95 %, para 12 semanas.	60
5.4	Estimación a posteriori de los momentos y cuantiles de los parámetros del modelo SVM-TVP con $\alpha_t = 0$ guardando 12 semanas.	61
5.5	Predicciones e intervalos de credibilidad, al nivel del 95 %, del modelo SVM-TVP con $\alpha_t = 0$ para 12 semanas.	64
5.6	Estimación de los parámetros y sus desviaciones estándar de los modelos GARCH guardando 12 semanas.	64
5.7	Estimación de los parámetros y sus desviaciones estándar de los modelos SV guardando 12 semanas.	65
5.8	Logaritmo de verosimilitud marginal, integral, BIC, AIC, FB, Q, Q2 y RECPM de los modelos: GARCH, SV, GARCH-2, SV-2, GARCH-J, SV-J, GARCH-M, SV-M, SVM-TVP y SVM-TVP con $\alpha_t = 0$	65
5.9	Estimación a posteriori de los momentos y cuantiles de los parámetros del modelo SVM-TVP guardando 24 observaciones.	66
5.10	Predicciones e intervalos de credibilidad, al nivel del 95 %, para 24 semanas.	70
5.11	Estimación a posteriori de los momentos y cuantiles de los parámetros del modelo SVM-TVP con $\alpha_t = 0$ guardando 24 obsevaciones.	70
5.12	Predicciones e intervalos de credibilidad, al nivel del 95 %, del modelo SVM-TVP con $\alpha_t = 0$ para 24 semanas.	72
5.13	Estimación de los parámetros y sus desviaciones estándar de los modelos GARCH guardando 24 observaciones.	73
5.14	Estimación de los parámetros y sus desviaciones estándar de los modelos SV guardando 24 observaciones.	73
5.15	Logaritmo de verosimilitud marginal, integral, BIC, AIC, FB, Q, Q2 y RECPM de los modelos: GARCH, SV, GARCH-2, SV-2, GARCH-J, SV-J, GARCH-M, SV-M, SVM-TVP y SVM-TVP con $\alpha_t = 0$	74

Lista de algoritmos

1	Algoritmo de Metropolis Hasting (MH)	28
2	Algoritmo de Gibbs	29
3	Algoritmo de aceptación y rechazo (AR)	30
4	Algoritmo de aceptación rechazo de Metropolis Hasting (ARMH)	31

Resumen

En el presente trabajo se expone el modelo de volatilidad estocástica en media, con parámetros que varían en el tiempo (SVM-TVP), introducido por Chan (2017), para modelar el precio de crudo Oriente en el Ecuador. Se presenta un enfoque basado en la inferencia bayesiana para estimar el modelo y se utilizan métodos de Monte Carlo con cadenas de Márkov, como son: el algoritmo de Metropolis-Hasting (MH), el muestreo de Gibbs y el algoritmo de aceptación y rechazo de Metropolis-Hasting (ARMH) de Chan (2017) para simular la distribución a posteriori de los parámetros del modelo. Se propone una aplicación del trabajo de Chib (1995), la cual se basa en el muestreo de Gibbs, para estimar la función de verosimilitud marginal del modelo SVM-TVP. Además, se compara el rendimiento del modelo SVM-TVP con los otros modelos de volatilidad condicional como: GARCH, GARCH-M, GARCH-J, GARCH-2 y sus contrapartes con volatilidad estocástica, utilizando: la función de verosimilitud marginal, el factor bayesiano y los criterios de información de Akaike (AIC), y Schwarz (BIC). También se muestra el impacto de la caída de los precios del petróleo en el Ecuador, especialmente en la balanza comercial y el sector público. Por último, se plantean las posibles causas de la alta volatilidad de los precios del crudo.

Palabras clave: Modelo de volatilidad estocástica en media, con parámetros que varían en el tiempo (SVM-TVP); Inferencia bayesiana; Métodos de Monte Carlo; Algoritmo de Metropolis-Hasting (MH); Muestreo de Gibbs; Algoritmo de aceptación y rechazo de Metropolis-Hasting (ARMH); Factor bayesiano; Verosimilitud marginal; Modelos: GARCH, SV y SVM-TVP.

Abstract

The present work exposes the stochastic volatility in mean model, with time-varying parameters (SVM-TVP), proposed by Chan (2017), to model Oriente crude oil price from Ecuador. An approach based on Bayesian inference is presented to estimate the model and Markov chains Monte Carlo (MCMC) methods are used, such as Metropolis-Hasting (MH) algorithm, the Gibbs sampling and the acceptance-rejection Metropolis-Hasting (ARMH) algorithm described by Chan (2017) to simulate the a posteriori distribution of the parameters of the model. An application from Chib (1995) is proposed, which is based on the Gibbs sampling to estimate the marginal likelihood of the SVM-TVP model. In addition, the performance of the SVM-TVP model is compared with the other conditional volatility models such as GARCH, GARCH-M, GARCH-J, GARCH-2 and their counterparts with stochastic volatility, using marginal likelihood, the Bayesian factor, Akaike information criterion (AIC) and Schwarz information criterion (BIC). Also, the impact of the fall in oil prices in Ecuador is presented, especially in the trade balance and the public sector. Finally, the possible causes of the high volatility of oil prices are presented.

Key words: Stochastic volatility in mean model with time-varying parameters (SVM-TVP); Bayesian Inference; Markov chain Monte Carlos; Hastings-Metropolis (MH) algorithm; Gibbs sampling; Acceptance-rejection Metropolis-Hastings (ARMH) algorithm; Bayesian factor; Marginal likelihood; Models: GARCH, SV y SVM-TVP.

CAPÍTULO 1

Introducción

En este capítulo se presenta una breve introducción a los temas que se desarrollarán dentro de este trabajo de titulación. En la primera parte se presenta la importancia de los precios del petróleo para el Ecuador y las diferentes alternativas para modelar su alta volatilidad. En las próximas secciones se presenta una introducción a: los espacios de estados, los modelos autoregresivos de heterocedasticidad condicional y los modelos de volatilidad estocástica. Por último, se hace una introducción a la inferencia bayesiana, la cual servirá más tarde para estimar el modelo SVM-TVP.

1.1 Motivación

Desde finales del año 2014, la súbita caída del precio del petróleo ha afectado a la economía de los países exportadores, entre ellos el Ecuador donde el sector petrolero significó el 11 % del PIB en el 2014. Especialmente, en Ecuador la caída de los precios del crudo redujo las exportaciones de 13.275 a 6.660 millones de dólares entre el 2014 y 2015, respectivamente, lo cual produjo un crecimiento del déficit en la balanza comercial de 712 a 2.094 millones de dólares entre el 2014 y 2015. En el sector fiscal esto provocó una reducción de ingresos; el estado para cubrir este déficit aumento su deuda total (externa y pública) de 24.099 a 33.930 millones de dólares desde el 2014 al 2016, lo que deviene de suma utilidad para un país petrolero poder predecir los precios del crudo y su volatilidad para de esta manera estar mejor preparado para una futura crisis.

La varianza de los precios del petróleo ha tendido a aumentar en los últimos dos años, lo que ha producido malestar en las principales bolsas mundiales y países productores. La primera forma de modelar esta característica de los datos es el modelo autoregresivo de heterocedasticidad condicional (GARCH), desarrollado por Bollerslev (1986) para extender el trabajo previo sobre los modelos ARCH de Engle (1982). La segunda clase de modelos son los de volatilidad estocástica, desarrollados por Taylor (1994); en estos la volatilidad es una variable latente que sigue un proceso estocástico.

Koopman y Uspensky introducen el modelo de volatilidad estocástica en media (SV-M)

en Koopman (2002) como una alternativa del modelo GARCH-M de Engle et al. (1987), bajo el cual la volatilidad estocástica entra en la ecuación de observación como una covariable. Más recientemente, el modelo de volatilidad estocástica en media, con parámetros que cambian a través del tiempo en la media condicional, generaliza el modelo de volatilidad estocástica en media (SV-M).

1.2 Modelo de espacios de estados generalizado

Los modelos de espacios de estado generalizados son flexibles y adecuados para una variedad de datos longitudinales, lo cual ha sido demostrado en la literatura como Kitagawa (1987) y Jørgensen et al. (1999). Un paso crucial en la aplicación de todos estos modelos al análisis de datos es estimar la distribución condicional de las variables de estado dada la totalidad o parte de los datos observados, así como los otros parámetros desconocidos del modelo. De forma abstracta, no existe dificultad en derivar los estimadores de máxima verosimilitud (MLE); pero la implementación real es extremadamente desafiante debido a la falta de expresiones de forma cerrada para integrales típicamente de muy alta dimensionalidad en la función de verosimilitud. Por lo tanto, uno tiene que depender de varias aproximaciones que esencialmente se dividen en tres categorías: analítica, Monte Carlo, y numérica.

Un modelo de espacios de estados generalizado consiste de dos procesos estocásticos: un proceso de observación d -dimensional y_t y un proceso de estados q -dimensional x_t , como se expresa a continuación:

- El proceso de estados x_0, x_1, \dots , es una cadena de Márkov con condición inicial $x_0 \sim p_0(x)$ y distribución de transición condicional

$$x_t | x_{t-1} \sim g_t(x | x_{t-1}).$$

- El proceso de observación y_t es condicionalmente independiente dado el proceso de estado $\{x_t, t \geq 0\}$ y cada y_t es condicionalmente independiente de $x_s, s \neq t$. Dado x_t , la distribución condicional es

$$y_t | x_t \sim f_t(y_t | x_t).$$

El problema más importante en el modelamiento de espacios de estados es la estimación de los estados x_t de las observaciones. Muchos problemas en series de tiempo, tales como: estimación de parámetros, cálculo de la función de verosimilitud y la predicción de una serie de tiempo estacionaria o no estacionaria se pueden manejar al estimar los estados. El problema de la estimación de los estados se puede formular como una evaluación de la función de densidad de probabilidades (f.d.p.) condicional de $p(x_n | Y_t)$, donde: Y_t es el conjunto de observaciones $\{y_1, \dots, y_t\}$. Al corresponder a los tres diferentes casos: $n > t$, $n = t$ y $n < t$; el problema de la estimación se puede clasificar dentro de tres correspondientes categorías donde la f.d.p. condicional $p(x_n | Y_t)$ se conoce como: predictor, filtro y suavizador, respectivamente.

1.2.1 El modelo lineal gaussiano de espacio de estados

El modelo lineal gaussiano de espacio de estados:

$$y_t = Z_t \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, H_t), \quad (1.2.1.1)$$

$$\alpha_{t+1} = T_t \alpha_t + \eta_t \quad \eta_t \sim \mathcal{N}(0, Q_t), \quad (1.2.1.2)$$

donde: y_t es un vector $p \times 1$ llamado vector de observaciones o medidas, α_t es un vector $m \times 1$, denominado vector de componentes no observadas. Z_t y T_t son matrices de tamaños $p \times m$ y $m \times m$, respectivamente.

La idea esencial en que se fundamenta este modelo es que la evolución del sistema a través del tiempo está determinada por α_t de acuerdo a la segunda ecuación (1.2.1.2); pero a causa de que α_t no se puede observar directamente se debe basar el análisis en las observaciones y_t . La primera ecuación (1.2.1.1) se denomina ecuación de las observaciones y la segunda ecuación (1.2.1.2) de los estados o de transición.

Hipótesis del modelo:

- Las matrices Z_t , T_t , H_t , y Q_t inicialmente se asumen conocidas, y además, Z_t y T_{t-1} dependen de y_1, \dots, y_{t-1} .
- Los términos de error ε_t y η_t se suponen serialmente no correlacionados e independientes el uno del otro en todo momento del tiempo.
- Se asume que el estado inicial α_t sigue una distribución normal $\mathcal{N}(a_1, P_1)$ y es independientemente de los errores $\varepsilon_1, \dots, \varepsilon_n$ y η_1, \dots, η_n , donde: a_1 y P_1 se asumen conocidos.

La primera ecuación de (1.2.1.1) tiene la estructura de un modelo de regresión lineal donde el vector de coeficientes α_t , varía con el tiempo. La segunda ecuación (1.2.1.2) representa un modelo autoregresivo de primer orden.

Las matrices Z_t , T_t , H_t y Q_t se denominan hiperparámetros para distinguirlos de los elementos que intervienen en los vectores de estado, los cuales se pueden considerar como parámetros aleatorios.

Para el modelo lineal gaussiano de espacios de estados, las distribuciones de probabilidades de $x|x_{t-1}$ y $y_t|x_t$ son gaussianas, sus medias y matrices de varianza-covarianza se pueden obtener por medio del filtro de Kalman ¹.

Sin embargo, para modelos no lineales o no gaussianos de espacios de estados, se utilizan varias técnicas de aproximación. Por ejemplo, para modelos con funciones no lineales en las ecuaciones de observación y estados con error aditivo se pueden utilizar el filtro de Kalman extendido (ver página 195 de Anderson y Moore (1979)). Otro ejemplo, se observa en el

¹El filtro de Kalman es un procedimiento recursivo unificado de predicción y corrección que permite actualizar los parámetros del modelo en cada etapa de la estimación. Una descripción más detallada del filtro de Kalman se encuentra en la página 326 de Shumway y Stoffer (2011).

modelo básico de volatilidad estocástica de Harvey et al. (1994):

$$y_t = \beta \exp \frac{h_t}{2} \epsilon_t, \quad t \geq 1, \quad (1.2.1.3)$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \sigma_\eta \eta_t, \quad (1.2.1.4)$$

$$h_1 \sim \mathcal{N} \left(\mu, \frac{\sigma^2}{1 - \phi^2} \right), \quad (1.2.1.5)$$

donde: ϵ_t y η_t siguen la distribución normal y no están correlacionadas. Se puede transformar a un modelo lineal al tomar el logaritmo de los cuadrados de las observaciones:

$$\log y_t^2 = h_t + \log \epsilon_t^2.$$

Finalmente, se emplea el filtro de Kalman para estimar los parámetros $\theta = (\phi, \sigma_\eta^2, \mu) \in (-1, 1) \times \mathbb{R} \times \mathbb{R}$ al maximizar una función de quasi-verosimilitud (ver apéndice de Kim et al. (1998)), la cual supone que $\log \epsilon_t^2$ sigue una distribución normal, esta última es una pobre aproximación.

1.3 Modelando la volatilidad de los precios del crudo Oriente

En esta sección se introducen los modelos de volatilidad condicional y sus contra partes estocásticas, ambos formulados en forma de espacios de estados. Además, se introduce el modelo de volatilidad estocástica en media, con parámetros que varían en el tiempo (SVM-TVP) aparecido por primera vez en Chan (2017), el cual generaliza la especificación del modelo de volatilidad estocástica en media (SV-M) de Koopman (2002).

1.3.1 Modelos GARCH

En esta sección se describen los modelos auto-regresivos con heterocedasticidad condicional generalizados (GARCH), introducidos en Bollerslev (1986) para extender el trabajo anterior sobre los modelos ARCH, anteriormente expuestos en Engle (1982). Los modelos GARCH son ampliamente utilizados para modelar precios del crudo (ver Chan y Grant (2016)).

El primer modelo es GARCH (1,1):

$$y_t = \mu + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_t^2), \quad (1.3.1.1)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad (1.3.1.2)$$

donde: $\epsilon_0 = 0$ y σ_0^2 es una constante. Para asegurar que el proceso de la varianza σ_t^2 sea siempre positivo y estacionario, se asume que $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$ y $\alpha_1 + \beta_1 < 1$. Se observa que la varianza condicional σ_t^2 es una función determinista de los parámetros del modelo y de la varianza pasada. A este modelo se lo conoce simplemente como GARCH.

La varianza condicional σ_t^2 en (1.3.1.2) sigue un proceso AR(1). A continuación, se introduce el modelo GARCH(2,1), en el cual σ_t^2 sigue el proceso AR(2),

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2,$$

donde: $\sigma_{-1}^2 = \varepsilon_0 = 0$ y σ_0^2 es constante. Una vez más, para asegurar que la varianza del proceso σ_t^2 es siempre positiva y estacionaria, se asume que los parámetros α_0 , α_1 , β_1 y β_2 son todos positivos y $\alpha_1 + \beta_1 + \beta_2 < 1$. A este modelo se lo conoce como GARCH-2.

El tercer modelo GARCH permite la posibilidad de infrecuentes “saltos” en la serie de datos, los cuales se pueden adaptar a cambios drásticos en los precios del crudo. Más específicamente, se considera el siguiente modelo GARCH con saltos (GARCH-J), como se expresa a continuación:

$$y_t = \mu + k_t q_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_t^2),$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 (y_{t-1} - \mu)^2 + \beta_1 \sigma_{t-1}^2,$$

donde: $q_t \in \{0, 1\}$ es una variable de salto con probabilidad de éxito, con $\Pr(q_t = 1) = \kappa$. Por lo tanto, si $q_t = 1$, un salto ocurre en el tiempo t y su tamaño es determinado por k_t , el cual se modela como $\varepsilon_t \sim \mathcal{N}(\mu_k, \sigma_k^2)$.

Por último, se establece el modelo GARCH en media (GARCH-M), bajo el cual la varianza condicional tiene dentro de su expresión a la media condicional como una covariable, como:

$$y_t = \mu + \lambda \sigma_t^2 + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_t^2), \quad (1.3.1.3)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 (y_{t-1} - \mu - \lambda \sigma_{t-1}^2)^2 + \beta_1 \sigma_{t-1}^2. \quad (1.3.1.4)$$

Esta variante permite que la serie de tiempo dependa de su propia volatilidad. Se tiene que cuando $\lambda = 0$, el modelo GARCH-M se reduce al modelo GARCH.

1.3.2 Modelos con volatilidad estocástica

Estos modelos se mencionan por primera vez en Taylor (1994). La volatilidad bajo un modelo volatilidad estocástica (SV) es una variable aleatoria, en contraste al modelo GARCH, en el cual la varianza condicional es una función determinista de los parámetros del modelo y los datos pasados de la serie. A continuación, se introducen los cuatro modelos con volatilidad estocástica (SV) que son contrapartes muy cercanas de los modelos GARCH descritos en la sección anterior.

En primer lugar, el modelo de Volatilidad Estocástica estándar (SV) es

$$y_t = \mu + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, \exp^{h_t}), \quad (1.3.2.1)$$

$$h_t = \mu_h + \phi_h (h_{t-1} - \mu_h) + \varepsilon_t^h, \quad \varepsilon_t^h \sim \mathcal{N}(0, \omega_h^2), \quad (1.3.2.2)$$

donde: el logaritmo de la volatilidad h_t sigue un proceso estacionario AR(1), con: $|\phi_h| < 1$ y media no condicional μ_h . El proceso es inicializado con $h_1 \sim \mathcal{N}(\mu_h, \omega_h^2 / (1 - \phi_h^2))$.

En el segundo modelo de Volatilidad Estocástica, la ecuación de observación es una vez más (1.3.2.1); pero el logaritmo de la volatilidad h_t ahora sigue el siguiente proceso estacionario AR(2):

$$h_t = \mu_h + \phi_h (h_{t-1} - \mu_h) + \rho_h (h_{t-2} - \mu_h) + \varepsilon_t^h, \quad \varepsilon_t^h \sim \mathcal{N}(0, \omega_h^2),$$

donde se asume que las raíces del polinomio característico asociadas con (ϕ_h, ρ_h) se encuentran fuera del círculo unitario. Además, se supone que h_1 y h_2 siguen una distribución no condicional:

$$h_1, h_2 \sim \mathcal{N} \left(\mu_h, \frac{(1 - \rho_h) \omega_h^2}{(1 + \rho_h) ((1 - \rho_h)^2 - \phi_h^2)} \right).$$

Este modelo de volatilidad estocástica se lo conoce como SV-2, el cual se reduce al modelo SV estándar cuando $\rho_h = 0$.

Similar al modelo GARCH-J, el tercer modelo de Volatilidad Estocástica aloja la posibilidad de saltos poco frecuentes. Específicamente, bajo el modelo de volatilidad estocástica con saltos (SV-J), la ecuación de observación es

$$y_t = \mu + k_t q_t + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, \exp h_t),$$

donde el logaritmo de la volatilidad h_t sigue el mismo proceso AR(1) como en (1.3.2.2). El indicador de salto q_t y el tamaño de salto k_t se modelan exactamente como en el GARCH-J.

Finalmente, se introduce el modelo de volatilidad estocástica en media (SV-M) de Koopman (2002), en el cual la volatilidad estocástica entra en la ecuación de observación como covariable:

$$y_t = \mu + \lambda \exp h_t + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, \exp h_t).$$

El logaritmo de la volatilidad sigue el mismo proceso AR(1) como en (1.3.2.2). El parámetro λ captura el grado de retroalimentación de la volatilidad. El modelo SV-M se reduce al modelo SV estándar cuando $\lambda = 0$.

Se resumen los modelos GARCH y SV en el Cuadro 1.1. Ambos modelos: GARCH y SV se estiman utilizando técnicas bayesianas. La estimación se describe en Chan y Grant (2016).

Cuadro 1.1: Lista de modelos GARCH y volatilidad estocástica

Modelos GARCH	
GARCH	GARCH(1,1) donde σ_t^2 sigue un proceso estacionario AR(1).
GARCH-2	Igual a GARCH, pero σ_t^2 sigue un proceso estacionario AR(2).
GARCH-J	Igual a GARCH, pero la ecuación de observación tiene una componente de “salto”.
GARCH-M	Igual a GARCH, pero σ_t^2 entra en la ecuación de observación como una covariable.
Modelos de volatilidad estocástica	
SV	Modelo de volatilidad estocástica, donde h_t sigue un proceso estacionario AR(1).
SV-2	Igual a SV, pero el log de la volatilidad h_t sigue un proceso estacionario AR(2).
SV-J	Igual a SV, pero la ecuación de observación tiene una componente de “salto”.
SV-M	Igual a SV, pero h_t es una covariable en la ecuación de observación.

1.3.3 Modelo de volatilidad estocástica en media con parámetros que varían en el tiempo

Se presenta el modelo de volatilidad estocástica, con parámetros que varían a través del tiempo (SVM-TVP), introducido por primera vez en Chan (2017), donde la volatilidad

estocástica también entra en la ecuación de la media condicional, como se expresa a continuación:

$$y_t = \tau_t + \alpha_t \exp h_t + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, \exp h_t), \quad (1.3.3.1)$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \beta y_{t-1} + \varepsilon_t^h, \quad \varepsilon_t^h \sim \mathcal{N}(0, \sigma^2), \quad (1.3.3.2)$$

$$\gamma_t = \gamma_{t-1} + \varepsilon_t^\gamma, \quad \varepsilon_t^\gamma \sim \mathcal{N}(0, \Omega), \quad (1.3.3.3)$$

donde y_t es la serie temporal de los precios del petróleo, el logaritmo de la volatilidad estocástica h_t sigue un proceso estacionario $AR(1)$ en (1.3.3.2), con $|\phi| < 1$ que depende de los precios pasados (y_{t-1}) y es inicializado con:

$$h_1 \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{1 - \phi^2}\right).$$

Además, el vector de coeficientes ($\gamma_t = (\alpha_t, \tau_t)'$), sigue un proceso de caminata aleatoria en (1.3.3.3), inicializado con $\gamma_1 \sim \mathcal{N}(\gamma_0, \Omega_0)$ y Ω es su matriz de covarianzas 2×2 . Finalmente, para completar la especificación del modelo, se asume que ε_t^y y ε_t^h no están mutuamente y serialmente correlacionados, y las distribuciones a priori para μ , ϕ , β , σ^2 y Ω son independientes, y están dadas por:

$$\begin{aligned} \mu &\sim \mathcal{N}(\mu_0, V_\mu), & \phi &\sim \mathcal{N}(\phi_0, V_\phi), & \beta &\sim \mathcal{N}(\beta_0, V_\beta), \\ \sigma^2 &\sim \mathcal{IG}(v_{\sigma^2}, S_{\sigma^2}), & \Omega &\sim \mathcal{IW}(v_\Omega, S_\Omega), \end{aligned} \quad (1.3.3.4)$$

donde $\mathcal{IG}(\cdot, \cdot)$ se refiere a la distribución Gamma Inversa y $\mathcal{IW}(\cdot, \cdot)$ representa a la distribución inversa de Wishart.

Las ecuaciones (1.3.3.1) y (1.3.3.2) generalizan el modelo original en Koopman (2002), al permitir que los parámetros de la media condicional de y_t puedan variar a través del tiempo, es decir, que τ_t y α_t estén variando en el tiempo. Además, se tiene que el modelo definido en (1.3.3.1) a (1.3.3.3), cuando $\alpha_t = 0$ para $t = 1, \dots, T$, se reduce al modelo estándar de volatilidad estocástica, con parámetros que varían en el tiempo (SV-TVP) o también conocido como SVM-TVP con $\alpha_t = 0$. Al permitir que $\alpha_t \neq 0$, el modelo se vuelve más resistente a cambios bruscos en h_{t-1} que afectarían a h_t , puesto que la volatilidad estocástica en escala logarítmica sigue un proceso $AR(1)$.

El modelo SVM-TVP dado por las ecuaciones (1.3.3.1) a (1.3.3.3) define un espacio de estados gaussiano con dos tipos de estados: γ_t y h_t (variables latentes), en el cual no es fácil estimar sus parámetros por métodos estándar de máxima verosimilitud, puesto que la ecuación de observación (1.3.3.1) no es lineal con respecto a la variable de estado h_t , por lo que no se puede aplicar directamente el filtro de Kalman².

1.3.3.1 Formulación del modelo SVM-TVP en estado de espacios

El modelo SVM-TVP, dado por las ecuaciones (1.3.3.1), (1.3.3.2) y (1.3.3.3), se puede re-formular como un espacio de estados. Al considerar (1.3.3.2) y (1.3.3.3), se tiene:

²Una descripción detallada del filtro de Kalman se encuentra en la página 326 de Shumway y Stoffer (2011).

$$\begin{bmatrix} h_{t+1} \\ \tau_{t+1} \\ \alpha_{t+1} \\ 1 \end{bmatrix} = \begin{bmatrix} \phi & 0 & 0 & \mu - \mu\phi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_t \\ \tau_t \\ \alpha_t \\ 1 \end{bmatrix} + \begin{bmatrix} \beta \\ 0 \\ 0 \\ 0 \end{bmatrix} y_t + \begin{bmatrix} \varepsilon_{t+1}^h \\ \varepsilon_{t+1}^\tau \\ \varepsilon_{t+1}^\alpha \\ 0 \end{bmatrix}. \quad (1.3.3.5)$$

Al establecer:

$$x_t = \begin{bmatrix} h_t \\ \tau_t \\ \alpha_t \\ 1 \end{bmatrix}, \quad \xi_t = \begin{bmatrix} \varepsilon_t^h \\ \varepsilon_t^\tau \\ \varepsilon_t^\alpha \\ 0 \end{bmatrix},$$

$$A = \begin{bmatrix} \phi & 0 & 0 & \mu - \mu\phi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} \beta \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

se puede reescribir la ecuación de estados (1.3.3.5), como se expresa a continuación:

$$x_{t+1} = Ax_t + By_t + \omega_{t+1},$$

donde y_t entra en la ecuación de estados como una variable exógena. Por último, faltaría la ecuación de observación para especificar el espacio de estados, la cual está dada por:

$$y_t = \tau_t + \alpha_t \exp h_t + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, \exp h_t).$$

El modelo SVM-TVP no es fácil de estimar sus parámetros por métodos estándar de máxima verosimilitud, puesto que la ecuación de observación no es lineal con respecto a la variable de estado h_t , por lo que no se puede aplicar directamente el filtro de Kalman.

1.4 Inferencia bayesiana

La estadística bayesiana es la rama de la estadística que se concentra alrededor de la fórmula de Bayes:

$$p(y|x) = \frac{p(x|y)p(y)}{\int p(x|y)p(y)dy} \propto p(x|y)p(y). \quad (1.4.0.6)$$

(El símbolo matemático ‘ \propto ’ se utiliza para indicar que dos valores son proporcionales). Para apreciar por completo la inferencia bayesiana, es importante entender que su razonamiento es diferente al de la estadística clásica. En la estadística clásica, los parámetros del modelo son desconocidos pero fijos; en cambio, en la estadística bayesiana, los parámetros de los modelos son tratados como aleatorios en vez de cantidades fijas. Se utilizará la siguiente notación:

1. Las funciones de densidades de probabilidades (f.d.p.) y las f.d.p. condicionales siempre utilizarán la misma letra p . Por ejemplo, en vez de escribir $f_X(x)$ y $f_Y(y)$

para las f.d.p. de X y Y , uno simplemente escribe $p(x)$ y $p(y)$. Similarmente, la f.d.p. condicional de X dado Y $f_{X|Y}(x|y)$ se denota como $p(x|y)$. Este estilo de notación es de gran valor descriptivo, a pesar de su aparente ambigüedad y se utilizará a lo largo de este trabajo de titulación.

2. No se hace distinción entre variables aleatorias y sus posibles valores. Ambas son usualmente indicadas en letras minúsculas. Se asume que el contexto está claro si una variable se puede interpretar como el posible valor que toma una variable (número) o una variable aleatoria.

El enfoque general de la estadística bayesiana está dado, como se sigue a continuación (comparado al enfoque clásico). Se asume que el vector de datos x ha sido obtenido a partir de una f.d.p. condicional $p(x|\theta)$, donde θ es un vector aleatorio de parámetros. La f.d.p. de θ transmite una información a priori (existente de antemano o antes de alguna experiencia) acerca de θ . Al observar los datos x se afecta el conocimiento de θ y la forma para actualizar esa información es utilizar la fórmula de Bayes (1.4.0.6). Los conceptos generales están resumidos en la siguiente definición.

Definición 1 (A priori, verosimilitud y a posteriori). *Sean x y θ los datos y parámetros en un modelo estadístico bayesiano.*

- *La f.d.p. de θ se llama la f.d.p. a priori.*
- *La f.d.p. condicional $p(x|\theta)$ se llama la función de verosimilitud bayesiana.*
- *El objeto central de interés es la f.d.p. $p(\theta|x)$, el cual por el teorema de Bayes, es proporcional al producto de la f.d.p. a priori y la función de verosimilitud:*

$$p(\theta|x) \propto p(x|\theta)p(\theta).$$

La f.d.p. a posteriori transmite el conocimiento de θ después de tener en cuenta la información de x . También nótese que la función de verosimilitud en la estadística bayesiana difiere de la estadística clásica. En la estadística bayesiana, la función de verosimilitud $p(x|\theta)$ es una f.d.p. condicional de los datos x , mientras en la estadística clásica, la función de verosimilitud $L(\theta; x) = p(x; \theta)$ es vista como una función de θ para una x fija. La f.d.p. a posteriori se puede ver como una versión a escala de la verosimilitud clásica. En efecto, si la f.d.p. a priori es constante, entonces la f.d.p. a posteriori coincide con la función de verosimilitud clásica, multiplicada por una constante.

1.4.1 Estimador bayesiano

En la teoría de la estimación y la teoría de la decisión, un estimador de Bayes o una acción de Bayes es un estimador o regla de decisión que minimiza el valor esperado posterior de una función de pérdida, es decir, la pérdida esperada a posteriori.

Se supone que un parámetro desconocido θ tiene una distribución a priori π . Sean $\hat{\theta} = \hat{\theta}(x)$ un estimador de θ , y $L(\theta, \hat{\theta})$ una función de pérdida o riesgo, como el error cuadrático

medio. El riesgo de Bayes de $\hat{\theta}$ se define como:

$$E_{\pi} \left\{ L(\theta, \hat{\theta}) \right\},$$

($E \{.\}$ denota la esperanza matemática) donde la esperanza se toma sobre la distribución de probabilidad de θ , de esta manera, la función de riesgo define una función de θ . Se dice que un estimador $\hat{\theta}$ es un estimador de Bayes si minimiza el riesgo de Bayes entre todos los estimadores. Equivalentemente, el estimador que minimiza la pérdida esperada a posteriori:

$$E \left\{ L(\theta, \hat{\theta}) | x \right\},$$

para cada x también minimiza el riesgo de Bayes y por lo tanto es un estimador de Bayes.

El error cuadrático medio (ECM) es la función de riesgo más comúnmente utilizada, la cual está definida por:

$$\text{ECM} = E \left\{ \left(\hat{\theta}(x) - \theta \right)^2 \right\},$$

donde la esperanza se toma sobre la distribución conjunta de θ y x . Al utilizar el ECM como función de riesgo, la estimación de Bayes del parámetro desconocido es simplemente la media de la distribución a posteriori (Teorema 1.1 de la página 228 de Lehmann y Casella (1998)):

$$\hat{\theta}(x) = E \{ \theta | x \}.$$

Este se lo conoce como el estimador de mínimo error cuadrático medio (MECM).

Parte I

Marco Teórico

CAPÍTULO 2

Historia e impacto del petróleo en la economía del Ecuador y su incertidumbre

En el segundo capítulo, se aborda en un rápido repaso el impacto del petróleo en el Ecuador, en especial las consecuencias en la economía de la caída de sus precios a finales del año 2014 en los sectores: fiscal, no financiero, balanza comercial y en su producción. En la primera sección se muestra una reseña histórica del descubrimiento del petróleo en el Ecuador y el inicio de su extracción, a continuación, se describe la relación histórica del petróleo con la economía ecuatoriana desde su descubrimiento a principios de la década de los 70. Más tarde, se muestra el impacto de los precios del crudo en la balanza comercial, PIB y sector fiscal en los últimos años. Finalmente, se enumera los factores externos e internos al sector petrolero que influyen a la volatilidad de los precios del crudo.

2.1 Descubrimiento del petróleo en Ecuador

La empresa Anglo perforó el primer pozo petrolero en el Ecuador, en Ancón, península de Santa Elena, el año de 1911. Sin embargo, la producción a niveles comerciales no comenzó hasta 1925 y la exportación inició en 1928, aunque en cantidades marginales. Hasta 1971, las exportaciones petroleras no superaban el 6 % del total de las exportaciones totales del Ecuador, según datos del Banco Central.

Entre 1928 y 1957, el país exportó 42 millones de barriles de crudo, igual al volumen exportado solo en 1972, año en que se inaugura la era del ‘boom’ petrolero. Durante cerca de cuarenta años, desde 1928 hasta 1959, la explotación de crudo se concentró en la península de Santa Elena. Sin embargo, en esos años varias compañías extranjeras: Shell, Standar Oil, California Oil, Tennessee, y la Western Geophysical Co obtuvieron más de 5 millones de hectáreas en nuevas concesiones para realizar exploraciones petroleras tanto en el litoral ecuatoriano como en la región Amazónica. Ello lo relata el libro PetroEcuador (2005).

El descubrimiento de nuevos pozos petroleros en la región amazónica después de 1967 convirtió al país en un productor mundial de petróleo y conllevó al aumento de los ingresos del gobierno de Velasco Ibarra al principio de 1972. Este año se terminó el Oleoducto

Transecuatoriano, un oleoducto de 503 kilómetros desde el Oriente a la ciudad portuaria de Esmeraldas. También se construyó una refinería al sur de la ciudad de Esmeraldas. Gracias a las exportaciones de crudo, las ganancias netas de divisas aumentaron de 43 millones de dólares en 1971 a 370 millones de dólares en 1974 (ver Flores y Merrill (1989)).

2.2 Relación histórica entre sector petrolero y la economía ecuatoriana

La producción y exportación de petróleo, que se inició a principios de la década de los 70 junto con dramáticos aumentos de precios internacionales del petróleo debido al bloqueo de la OPEP, contribuyeron significativamente a un crecimiento económico sin precedentes. El Producto Interno Bruto aumentó en promedio en más del 9 % por año entre 1970 y 1977 en comparación con sólo el 5,9 % de 1960 a 1970. El PIB real del sector manufacturero solo experimentó una tasa de crecimiento anual promedio del 12,9 % del PIB real durante 1975-1977. Ecuador se convirtió en un país de ingresos medio bajos; a pesar de esto, Ecuador continuó siendo uno de los países más pobres de América del Sur Flores y Merrill (1989).

Las importaciones reales aumentaron en un promedio anual de 7 % entre 1974 y 1979; esto dio lugar a un patrón de inflación que erosionó los ingresos. Durante el mismo período, la deuda externa del país pasó de 324 millones a cerca de 4,5 mil millones de dólares (ver Flores y Merrill (1989)).

En el año 1985 Ecuador se retiró por un año de la Organización de Países Exportadores de Petróleo (OPEP) con el fin de liberarse de las cuotas de exportación de esa organización y por lo tanto aumentar los ingresos de exportación de petróleo. En 1984 el petróleo representó aproximadamente el 70 % de todas las exportaciones de productos básicos y alrededor del 50 % de los ingresos del gobierno central. En 1985 el Ecuador obtuvo más de 1,8 mil millones de dólares en ingresos procedentes de las exportaciones de petróleo, dos tercios de los ingresos por exportaciones de Ecuador ese año.

Sin embargo, una fuerte caída de los precios internacionales del petróleo en 1986 dio lugar a una caída de las exportaciones a 1,1 miles de millones de dólares en ingresos por exportación de petróleo. La cuenta corriente de la balanza de pagos, que registró un superávit de 149 millones de dólares en 1985, mostró un déficit de 613 millones de dólares para 1986. Las reservas sobre tipos de cambio se redujeron a 145 millones de dólares para mediados de 1986, y el crecimiento real del PIB para 1986 llegó a sólo el 1,7 %, en comparación con el 3,8 % en 1985. Para hacer frente a la crisis económica, en enero de 1987, el gobierno suspendió los pagos de deuda a todas las instituciones de crédito privadas e impuso un recargo del 25 % en muchos artículos importados.

En marzo de 1987 un terremoto destruyó unos cuarenta kilómetros del oleoducto transecuatoriano y sus estaciones de bombeo, provocando una suspensión de casi seis meses en la producción del crudo y la pérdida de otros 700 millones de dólares en ingresos por exportaciones. El PIB cayó el 5,2 % en 1987, la inflación subió un 32,5 % y el déficit comercial se situó en 33 millones de dólares. El gobierno respondió a la emergencia financiera al elevar los precios internos de la gasolina en 80 %, y autobús y taxi en un 14 %. Para ayudar a compensar el déficit de ingresos del petróleo, un consorcio de bancos

internacionales prestó a Ecuador un adicional de 220 millones de dólares, con lo que la deuda externa del sector público a finales de 1987 llegó a ser de 9,6 mil millones de dólares, una de las de más altas del mundo per cápita (ver Flores y Merrill (1989)).

El deterioro de los resultados económicos de 1997 a 1998 culminó en una grave crisis financiera en 1999. La crisis se precipitó por una serie de choques externos, tales como: el fenómeno meteorológico de El Niño en 1997 y la fuerte caída de los precios mundiales del petróleo entre 1997 y 1998. Estos factores resaltaron la insostenible combinación de políticas económicas del Ecuador en grandes déficits fiscales y dio lugar a una contracción del 7,3 % del PIB, la inflación anual del año en 52,2 % y una devaluación del 65 % de la moneda nacional en 1999.

Impulsada por los precios más altos del petróleo, la economía ecuatoriana experimentó una recuperación modesta entre el 2000 y 2001, con un incremento del PIB del 2,3 % en 2000 y 5,4 % en 2001. El crecimiento del PIB se estabilizó a 3,3 % en 2002.

Ecuador recibió 77.530 millones de dólares por exportaciones petroleras entre el 2007 y el 2013, con precio promedio del crudo de 74,09 dólares, y con un precio máximo por barril de 140,1 dólares en julio de 2008 y mínimo de 23,8 dólares en marzo de 2009.

(a) Serie z_t (precios históricos del crudo Oriente)

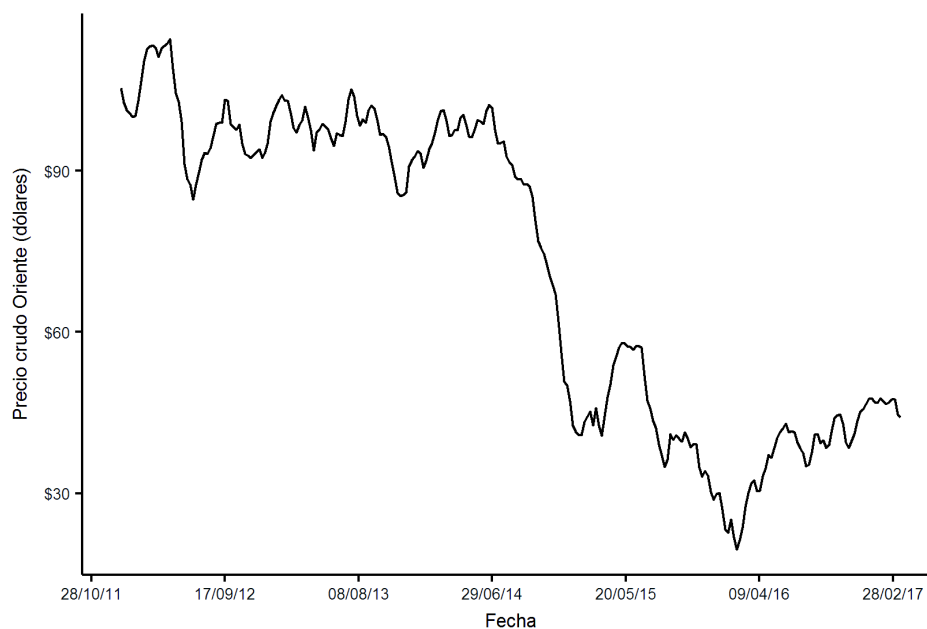


Figura 2.1 Evolución de los precios del crudo ecuatoriano.

Fuente: Sistema Nacional de Información (SNI)

Pese a que Ecuador recibió ingresos importantes por un precio alto, la lógica cambió en el 2014. Tras varios meses de caída, principalmente por el aumento de oferta de Estados Unidos del petróleo de esquisto (que se obtiene de rocas a través de técnicas de fracturación hidráulica); en noviembre la decisión de la Organización de Países Exportadores de Petróleo (OPEP) de no reducir su producción precipitó la caída.

2.3 Producción Petrolera del Ecuador

En la actualidad Ecuador extrae principalmente el crudo Oriente con 23 grados API¹ promedio, que es catalogado como crudo intermedio, con una producción de 6,8 millones de barriles en el mes de diciembre de 2017. Además, se extrae el crudo Napo de 18 a 21 grados API, catalogado como crudo pesado y agrio con una producción de 1,7 millones de barriles en el mismo mes, también cabe recalcar que la producción de este crudo se redujo en un 41,5 % de octubre a diciembre de 2017. En diciembre de 2017, las exportaciones de totales crudo fueron de 10,3 millones de barriles, 14,1 % menos que igual mes de 2016, como se puede ver en la Figura 2.2.

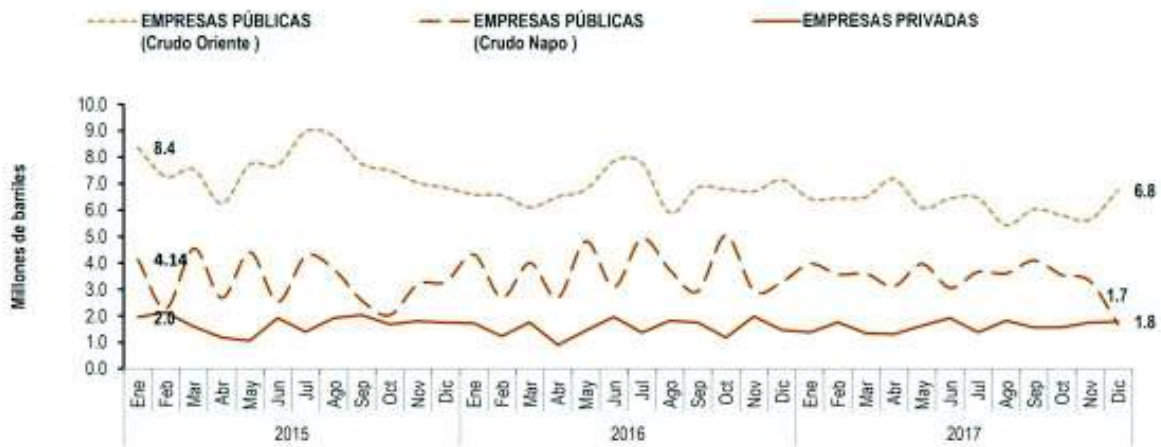


Figura 2.2 Evolución mensual de la producción petrolera ecuatoriana.

Fuente: Banco Central del Ecuador

En 2015 el Ecuador produjo 198,2 millones de barriles, 2,4 % menos que la producción de 2014. Las empresas, Petroamazonas EP y la operadora Río Napo, redujeron su producción en 3,6 % y 2,8 %, respectivamente; las compañías privadas también mostraron una baja en 6,9 % (ver Figura 2.3). La disminución de la producción, consecuencia de los precios bajos del petróleo, llevó al Gobierno Nacional a modificar las inversiones en este sector y a recortar la producción. El Gobierno Nacional priorizó los campos más rentables o con menores costos de producción.

La producción nacional presentó una contracción anual de 3,4 % en el 2017 con respecto al 2016. Esta caída se relaciona con el cumplimiento del acuerdo firmado por Ecuador en noviembre de 2016, en el cual se comprometió con la OPEP a reducir su producción a partir de enero de 2017. El análisis por empresas evidencia que Petroamazonas EP aumentó su producción en 6,6 % en 2017, debido a que el campo Sacha pasó a esta empresa desde agosto de 2016. Por su parte, las compañías privadas continuaron bajando su producción en 1,8 %.

¹Grados API: de sus siglas en inglés American Petroleum Institute, es una medida de densidad que, en comparación con el agua a temperaturas iguales, precisa cuán pesado o liviano es el petróleo. Índices superiores a 10 implican que son más livianos que el agua y, por lo tanto, flotarían en esta.

	Millones de Barriles				Tasa de Crecimiento		
	Enero - Diciembre				Enero - Diciembre		
	2014	2015	2016	2017	2015/2014	2016/2015	2017/2016
Producción Nacional Crudo en Campo	203.1	198.2	200.7	193.9	-2.4%	1.3%	-3.4%
Empresas Públicas	158.0	154.3	158.1	152.1	-2.3%	2.5%	-3.8%
Petroamazonas EP (B15)	131.8	127.3	142.7	152.1	-3.4%	12.1%	6.6%
Operadora Río Napo (Sacha)	26.2	27.0	15.4	0.0	3.2%	-42.8%	-100.0%
Compañías Privadas (excl. B15)	45.1	43.9	42.6	41.8	-2.6%	-3.0%	-1.8%

	Miles de Barriles			
	Enero - Diciembre			
	2014	2015	2016	2017
Producción Promedio Diaria Crudo en Campo	556	543	548	531
Empresas Públicas	433	423	432	417
Petroamazonas EP (B15)	361	349	390	417
Operadora Río Napo (Sacha)	72	74	42	0
Compañías Privadas (excl. B15)	124	120	116	115

Figura 2.3 Producción petrolera ecuatoriana por empresas.

Fuente: Banco Central del Ecuador

La producción diaria de crudo para las empresas públicas alcanzó 401,2 miles de barriles en promedio durante diciembre de 2017, lo cual representó una reducción del 6,85 % respecto al mismo mes en el año 2016 que alcanzó 430,7 miles de barriles en promedio (ver Figura 2.4). A diciembre del 2016, la producción de crudo en Ecuador se distribuía entre Petroamazonas EP y la Operadora Río Napo, el 82,5 % para Petroamazonas EP y 17,5 % para la Operadora Río Napo; en cambio, a diciembre de 2017 el 100 % de la producción le corresponde a Petroamazonas EP, luego de asumir la producción de la Operadora Río Napo.



Figura 2.4 Producción diaria promedio de Petroamazonas EP y Operadora Río Napo en miles de barriles.

Fuente: Banco Central del Ecuador

La producción diaria de crudo para las compañías privadas ha tendido a disminuir de 126,4 a 110 miles de barriles entre enero de 2015 a enero de 2017 (ver la Figura 2.5); pero durante el transcurso de 2017 aumentó hasta 118,6 miles de barriles registrados en diciembre de 2017.



Figura 2.5 Producción diaria promedio de las compañías privadas en miles de barriles.

Fuente: Banco Central del Ecuador

2.4 Precios históricos de los crudos Oriente, Napo y WTI

Entre 2008 y principios de 2011, se observa en la Figura 2.6 como el West Texas Intermediate (WTI) registró precios mayores a los precios de los crudos ecuatorianos: Oriente y Napo. Pero esta situación se revirtió para los crudos de origen ecuatoriano que se comercializaron a precios superiores al del WTI por varios meses, durante el periodo abril de 2011 a julio de 2013 para el caso del crudo Oriente y durante el periodo septiembre de 2011 a septiembre 2012 para el crudo Napo. Desde agosto de 2014, los precios internacionales del petróleo han mostrado una tendencia a la baja, determinados principalmente por un exceso de oferta en el mercado internacional. La tendencia a la baja terminó a mediados de 2016; aunque hasta la fecha el crudo no ha logrado alcanzar los precios comprendidos entre el 2011 y 2014.

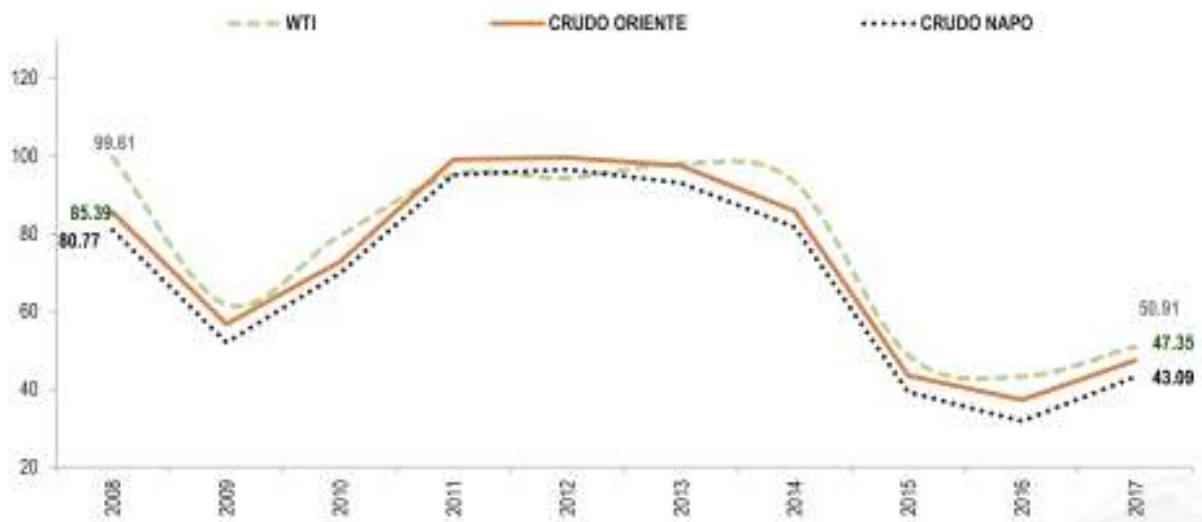


Figura 2.6 Evolución de los precios del crudo WTI, Oriente y Napo, en dólares.

Fuente: Banco Central del Ecuador

En la Figura 2.7 se puede apreciar que el West Texas Intermediate (WTI) registró un valor promedio de 57,9 dólares el barril en diciembre de 2017, lo cual muestra un incremento en su precio de 11,4 % con relación a similar mes de 2016. Los precios de los crudos Oriente y Napo alcanzaron un incremento anual de 26,2 % y 27,5 %, respectivamente, con relación a diciembre de 2016. Últimamente, desde julio de 2017 se ha incrementado la tendencia al alza del petróleo, debido a acuerdo de reducción de cuotas fijadas por la OPEP.

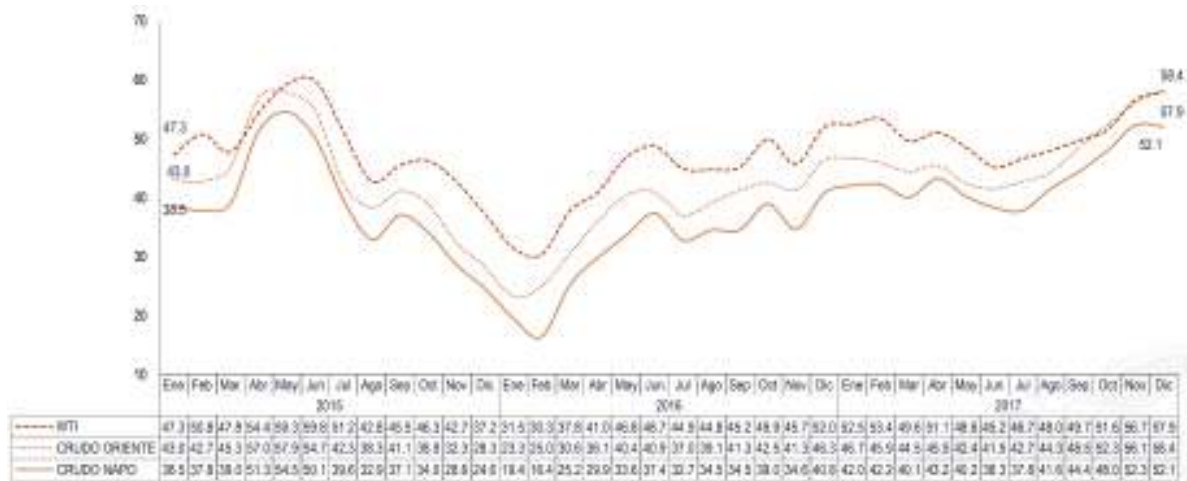


Figura 2.7 Evolución de los precios promedio mensuales del crudo WTI, Oriente y Napo en dólares.

Fuente: Banco Central del Ecuador

2.5 Impacto del precio del Petróleo

2.5.1 Impacto en la balanza comercial

La balanza comercial total registró en el año 2015 un déficit de 2.129,6 millones de dólares, lo cual significó un aumento en el déficit de 723,2 millones de dólares, 194 % en comparación al 2014. Este comportamiento se explica principalmente por una disminución de las exportaciones, debidas a la caída del precio del petróleo; en cambio, el año 2016 registró un superávit de 1.247 millones de dólares. Si se compara con el resultado del año 2015, se puede concluir que hubo una recuperación de la balanza comercial de 158,6 %, gracias al leve aumento de los precios del crudo que ha sucedido desde el mes de febrero de 2016 (ver Figura 2.8).

La balanza comercial total a diciembre de 2016 registró un superávit de 86,7 millones de dólares, lo cual significó una recuperación comercial de 308,9 % frente al saldo a noviembre de 2016 que registró un déficit de 41,5 millones de dólares (ver Figura 2.9).



Figura 2.8 Evolución anual de la balanza comercial petrolera y no petrolera frente a los precios promedios anuales del petróleo.

Fuente: Banco Central del Ecuador



Figura 2.9 Evolución de la balanza comercial mensual millones de dólares.

Fuente: Banco Central del Ecuador

Las exportaciones petroleras se redujeron dramáticamente de 13.275,9 a 6.660,3 millones de dólares entre el 2014 y 2015, lo cual significó que el Ecuador dejó de recibir 6.615,6 millones de dólares en exportaciones. Esta caída también produjo la reducción de la balanza comercial petrolera de 6.917,1 a 2.757,0 millones de dólares entre 2014 y 2015 (ver Figura 2.10). Este menor superávit se debe principalmente a una disminución en el valor unitario promedio del barril exportado de crudo, el cual pasó de 87,8 a 43,2 dólares entre el 2014 y 2015, respectivamente. Por su parte, la balanza comercial no petrolera disminuyó su déficit en 36 % frente al resultado contabilizado en el mismo período de 2014, es decir, se redujo el déficit de 7.640,2 a 4.886,6 millones de dólares entre el 2014 y 2015 (ver Figura 2.8).



Figura 2.10 Evolución anual de las exportaciones en millones de dólares.

Fuente: Banco Central del Ecuador

Entre enero y diciembre de 2016, las exportaciones totales en valor FOB² alcanzaron 16.797,7 millones de dólares, 8,4 % menor en términos relativos frente a las exportaciones realizadas en el mismo período de 2015 (ver Figura 2.10). En cambio, en 2017 las exportaciones totales en valor FOB alcanzaron 19.122,5 millones de dólares, lo cual representó 13,8 % más que las ventas externas realizadas en el mismo período de 2016.

2.5.2 Impacto en el sector fiscal

Debido a la caída de los precios del petróleo desde diciembre de 2014, el estado ecuatoriano ha dejado de recibir importantes ingresos que ha afectado a su tasa de crecimiento del PIB; además, ha impulsado al crecimiento de la deuda pública y externa para cubrir los huecos dejados por la falta de estos ingresos en el presupuesto general del Estado ecuatoriano.

2.5.2.1 Impacto en el sector público no financiero

El sector público no financiero (SPNF) comprende: presupuesto general del estado (PGE), empresas públicas no financieras (EPNF) y el conjunto de instituciones del resto del SPNF (RSPNF). Por concepto de ingresos del SPNF, el Ecuador tuvo una recaudación de 33.586 millones de dólares en 2015 (13,95 % menos que en el 2014 que fue de 39.032 millones de dólares). Al mismo tiempo, se registró una reducción del 12,7 % en los gastos del SPNF que en el 2014 fueron de 44.346 millones de dólares y se redujeron a 38.676 millones de dólares en el 2015. De esta manera, se generó un resultado global³ deficitario de

²Free on board: cláusula de comercio internacional en la cual el vendedor debe realizar el trámite para la exportación de la mercancía, así como asumir los costos de la misma.

³Resultado global: corresponde al resultado o déficit global y es aquel que se obtiene de la diferencia entre los ingresos (con incluir desembolsos) y los gastos (con incluir amortizaciones pero incluyendo el pago de intereses).

5.091 millones de dólares y un resultado primario⁴ deficitario de 3.723 millones de dólares en 2015.

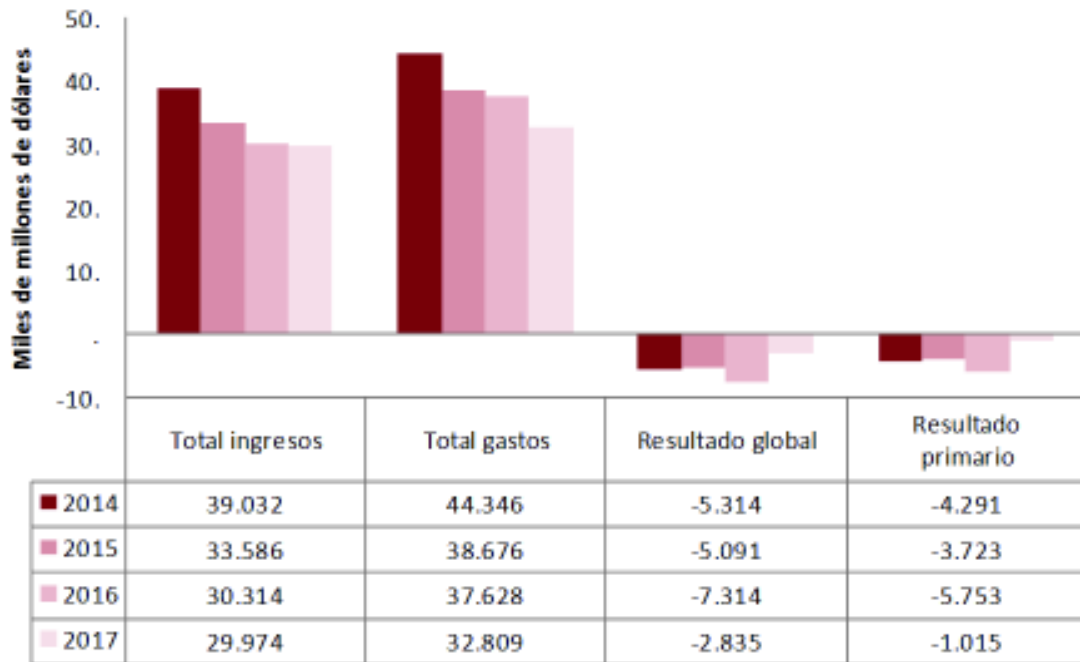


Figura 2.11 Evolución del sector público no financiero entre 2014 y 2016.

Fuente: Banco Central del Ecuador

Los ingresos petroleros fueron 6.346,2 millones de dólares en 2015, lo cual significó un reducción del 41,8 % en comparación al año 2014 que fue 10.905,8 millones de dólares. Los ingresos petroleros significaron el 10,8 % y 6,4 % en el 2014 y 2015, respectivamente. En la Figura 2.12 se muestra en detalle como los ingresos totales del Sector público no financiero (SPNF) decrecieron 14 % entre los años 2014 y 2015.

	Millones de dólares				% de crecimiento			% del P.I.B.			
	2014	2015	2016	2017	2014-2015	2015-2016	2016-2017	2014	2015	2016	2017
INGRESOS TOTALES	39.032,0	33.585,9	30.314,1	32.740,9	-14,0	-9,7	8,0	38,7	33,5	31,0	32,6
Petroleros	10.905,8	6.346,2	5.402,0	5.868,9	-41,8	-14,9	8,6	10,8	6,3	5,5	5,8
No Petroleros	23.939,2	25.942,0	24.294,3	25.287,2	8,4	-6,4	4,1	23,7	25,9	24,8	25,2
Ingresos tributarios	14.460,0	15.588,2	14.017,4	14.078,2	7,8	-10,1	0,4	14,3	15,6	14,3	14,0
IVA	6.375,6	6.352,3	5.399,7	5.978,8	-0,4	-15,0	10,7	6,3	6,3	5,5	6,0
ICE	803,2	839,6	789,6	936,8	4,5	-6,0	18,6	0,8	0,8	0,8	0,9
A la renta	4.160,7	4.734,2	3.639,7	3.763,9	13,8	-23,1	3,4	4,1	4,7	3,7	3,7
Arancelarios	1.357,1	2.025,7	1.632,7	1.467,9	49,3	-19,4	-10,1	1,3	2,0	1,7	1,5
Otros impuestos	1.763,3	1.636,3	2.555,8	1.930,8	-7,2	56,2	-24,5	1,7	1,6	2,6	1,9
Contribuciones Seguridad Social	4.718,1	5.055,4	4.741,4	5.414,8	7,2	-6,2	14,2	4,7	5,0	4,8	5,4
Otros	4.761,1	5.298,4	5.535,5	5.794,2	11,3	4,5	4,7	4,7	5,3	5,7	5,8
Resultado operacional emp. Púb. No financ.	4.187,0	1.297,6	617,9	1.584,8	-69,0	-52,4	156,5	4,1	1,3	0,6	1,6

Figura 2.12 Ingresos del SPNF entre 2014 y 2016.

Fuente: Banco Central del Ecuador

⁴Resultado primario o déficit primario consiste en la diferencia entre ingresos (con incluir desembolsos) y gastos (con incluir ni amortizaciones ni el pago de intereses).

2.5.2.2 Impacto en la deuda externa

A diciembre de 2015, el saldo de la deuda externa pública ascendió a 20.226 millones de dólares, lo cual significó un incremento del 15 % con respecto a diciembre de 2014 que fue de 17.582 millones de dólares. En cambio, a diciembre de 2015, la deuda externa privada ascendió a 7.535 millones lo que significó un incremento del 15,6 % con respecto a diciembre de 2014 que fue de 6.517 millones de dólares. Entre los años 2013 y 2015, la deuda externa total experimentó un incremento desenfrenado de 8.988 millones de dólares y el saldo de esta al 2015 fue de 27.761 millones de dólares, que representó el 27 % del PIB.

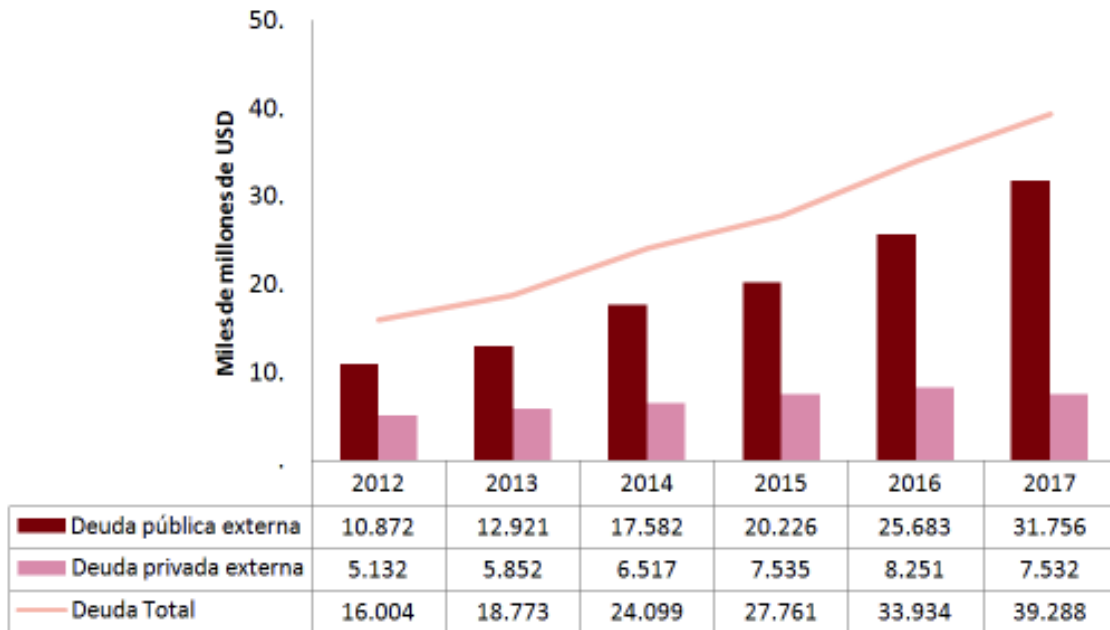


Figura 2.13 Evolución de la deuda externa (pública y privada) en millones de dólares.

Fuente: Banco Central del Ecuador

2.5.2.3 Impacto en la deuda interna

A diciembre de 2015, la deuda interna del Ecuador ascendía a 12.546 millones de dólares, lo cual representó el 12,66 % con relación al PIB. A diciembre de 2014, la deuda interna ascendía a 12.558 millones de dólares, lo cual representó el 12,44 % del PIB. Últimamente, en abril de 2016⁵ la deuda interna aumentó a 12.919 millones de dólares, se incluyó la deuda con el IESS de 7.000 millones de dólares.

En la Figura 2.14 se puede observar la evolución de la deuda pública total. Hasta el año 2011, la deuda pública se mantuvo estable, alcanzó su mínimo en 2009 (10.199 millones de dólares). Desde el 2012 comenzó la tendencia del crecimiento de la deuda pública; aunque la deuda todavía no representa 40 % del PIB (desde el año 2000 hasta el 2016). Se estimó que la deuda pública representó el 32,4 % del PIB a diciembre de 2016.

⁵A partir de mayo de 2016, el Banco Central ya no consideró la deuda con el IESS para calcular la deuda interna.

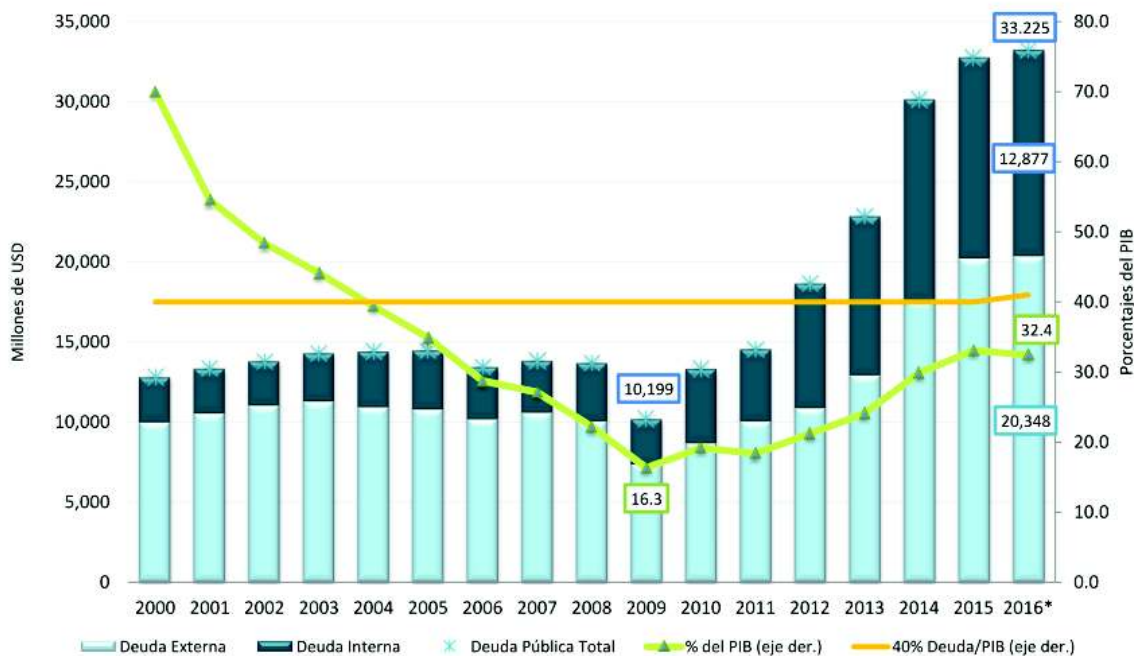


Figura 2.14 Evolución de la deuda externa e interna del Ecuador.

Fuente: Ministerio de Finanzas

2.6 Incertidumbre en los precios del petróleo

A diferencia de lo que se pensaría, los precios del petróleo no dependen solamente de la producción o cuotas diarias y costos de extracción. Además, existen factores exógenos a la industria petrolera muy importantes que producen incertidumbre en la volatilidad de los precios como, por ejemplo, los aspectos geopolíticos. Estos cambios se deben a eventos exógenos a la actividad económica; sin embargo, tienen efectos reales en la economía del mundo, especialmente los países productores como el Ecuador.

Los cambios en la volatilidad de los precios del petróleo se pueden generar por dos factores exógenos (ajenos) a la producción o mercado del petróleo, estos son: los eventos relacionados con los miembros de la OPEP y la incertidumbre macroeconómica en general.

En primer lugar, los eventos relacionados con los miembros de la OPEP puede deberse a guerras, tensiones diplomáticas, levantamientos políticos y crisis económicas que afectan a uno o más países miembros de la OPEP. Para lograr aislar estos eventos exógenos a la actividad petrolera, en Plante y Traum (2015) se construye un índice, basado en el conteo de artículos en los tres periódicos norteamericanos más importantes: the Wall Street Journal, the Financial Times, y the New York Times, utilizando la palabra clave OPEP. En la Figura 2.15 se muestra la evolución histórica del índice desde 1993 hasta 2016, por trimestres. En la Figura 2.15 se observa que hay un alto número de cúspides cercanas a fechas de eventos de la OPEP, en particular a reuniones de emergencia. Otras cúspides incluyen eventos como la Primavera Árabe, la huelga de productores en Venezuela y la invasión norteamericana a Iraq.

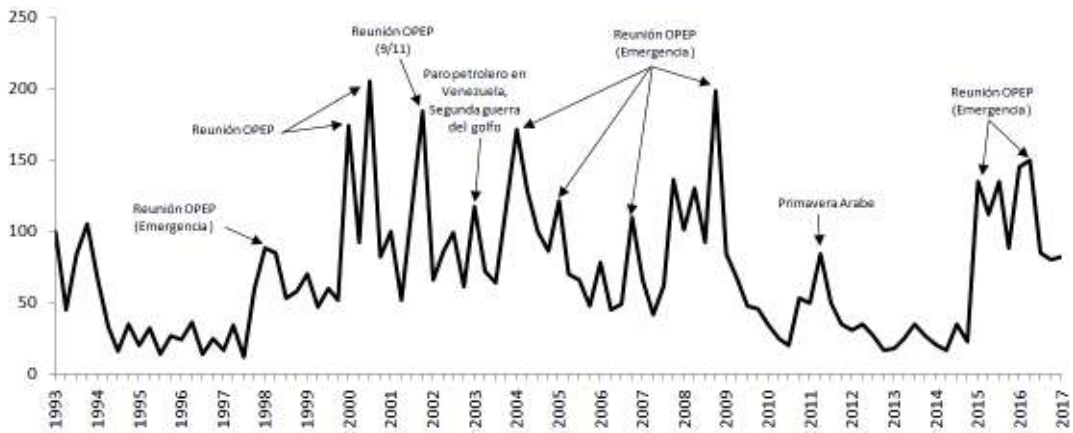


Figura 2.15 Evolución del índice de las noticias relacionadas con la OPEP.

Fuente:Plante y Traum (2015)

El segundo factor exógeno es la incertidumbre macroeconómica, la cual se puede medirla utilizando el índice VIX de la Chicago Board Operations Exchange (CBOE), el cual mide la expectativa a plazo cercano de la volatilidad en el índice S&P 500. En cambio, para medir la incertidumbre de los precios del petróleo se puede utilizar el índice bursátil OVX, el cual mide las expectativas del mercado a corto plazo en los precios nominales del crudo WTI. En la Figura 2.16 se gráfica el índice VIX versus OVX desde 2007 hasta 2016, por trimestres. Sin embargo, la Figura 2.16 sugiere una fuerte relación entre la incertidumbre económica y la incertidumbre de los precios del petróleo.

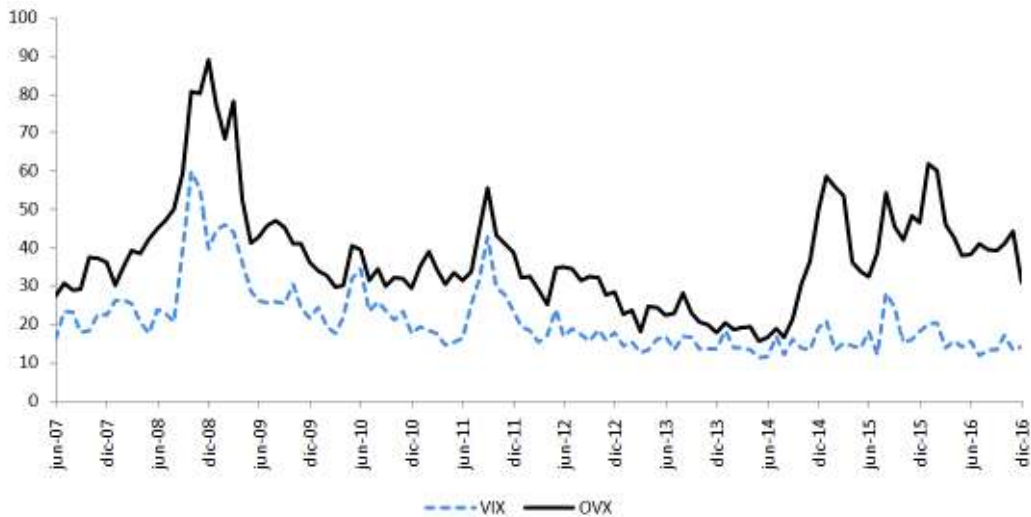


Figura 2.16 Medidas de incertidumbre:VIX (volatilidad del índice S&P 500) y OVX (volatilidad de los contratos futuros a un mes para WTI).

Fuente:Plante y Traum (2015)

Por ejemplo, ambos índices, VIX y OVX, crecen dramáticamente cuando la economía mundial entró en recesión al final de 2008 y los dos tienen una correlación cruzada de 0,77. Esta alta correlación entre VIX y OVX crea cierta incertidumbre acerca de anteriores fuentes que indicaban una relación negativa, entre la actividad económica de E.E.U.U. y

la incertidumbre del precio del petróleo, encontrada en trabajos empíricos previos. Esto se puede atribuir a eventos exógenos, por ejemplo, las convulsiones políticas en países de la OPEP. Alternativamente, podría ser una consecuencia de la incertidumbre del precio del petróleo producida por cambios en la incertidumbre macroeconómica, o quizás una combinación de ambas.

CAPÍTULO 3

Métodos de Monte Carlo basados en cadenas de Márkov (MCMC)

En este capítulo se describen el método de Monte Carlo de muestreo por aceptación y rechazo, y los basados en cadenas de Márkov (MCMC): el método de Metropolis-Hasting (MH), muestreo de Gibbs, y aceptación y rechazo de Metropolis-Hasting (ARMH). Más tarde, se los utilizarán para estimar el modelo SVM-TVP.

El objetivo de los MCMC es primero muestrear una función de distribución de probabilidad y entonces poder estimar esperanzas bajo la distribución simulada. Más específicamente, se quiere estimar la función de verosimilitud. La función de verosimilitud sobre los parámetros del modelo SVM-TVP, se define por una integral T-dimensional (dado los parámetros $\psi = (\mu, \phi, \beta, \sigma^2, \Omega)$):

$$L(\psi) = p(y|\psi) = \int p(y, h, \gamma|\psi) dh d\gamma \propto \int p(y|h, \gamma, \psi) p(h|\gamma, \psi) p(\gamma|\psi) dh d\gamma.^1$$

Al establecer $Y_{t-1} = (y_1, \dots, y_{t-1})$, la función de verosimilitud se expresa como una mezcla sobre las distribuciones de los estados:

$$L(\psi) = p(y|\psi) = \prod_{t=1}^T p(y_t|Y_{t-1}, \psi) = \prod_{t=1}^T \int p(y_t|h_t, \gamma_t, \psi) p(h_t, \gamma_t|\psi) dh_t d\gamma_t.$$

Esta función de verosimilitud es intratable, es decir, no se puede calcular fácilmente. La fuente del problema es que la f.d.p. $p(h_t, \gamma_t|\psi)$ no se puede expresar en forma cerrada; y además, $y_t|Y_{t-1}, \psi$ no tiene una expresión analítica. Como resultado, la función de verosimilitud se debe aproximar por un método de simulación. Esto también imposibilita el análisis directo de la f.d.p. a posteriori $p(\psi|y)$ (debido a que $p(\psi|y) \propto p(y|\psi)p(\psi)$). Una de las primeras soluciones planteadas fue el método de Aumento de datos propuesto por Tanner y Wong (1987). El algoritmo de Aumento de datos permite la estimación de los parámetros ψ , sin calcular la función de verosimilitud; básicamente, los parámetros de ψ son aumentados a las variables latentes h y γ para formar un vector de parámetros más grande.

¹Esta relación se deriva de la fórmula de Bayes. Sea A y B dos eventos, entonces el teorema de Bayes establece $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$ donde $P(A)$ y $P(B)$ son probabilidades marginales de eventos aleatorios de A y B , respectivamente. El factor de proporcionalidad es $P(B)$.

La f.d.p. conjunta de (h, γ, ψ) sobre y , se denota por $p(h, \gamma, \psi|y)$ (f.d.p. a posteriori); esta última, la f.d.p. a posteriori, es ahora el centro de atención.

El método de MCMC es simple y extremadamente general. Con el fin de obtener una muestra² de una distribución de probabilidad referida como la distribución objetivo (para el modelo SVM-TVP sería $p(h, \gamma, \psi|y)$), se construye una cadena de Márkov ergódica (irreducible³ y aperiódica⁴) reversible, puesto que la condición de la reversibilidad⁵ es una condición suficiente para que la n -ésima iteración converja a la distribución invariante que es la distribución objetivo (Teorema expuesto en Chib y Greenberg (1995)). Una vez construida la cadena de Márkov, una muestra de la distribución objetivo se obtiene al simular la cadena una gran cantidad de veces y guardando sus salidas.

3.1 Algoritmo de Metropolis-Hasting

Se supone que se tiene una f.d.p. candidata que puede generar a la f.d.p. objetivo ($f(x)$); la f.d.p. candidata generadora es $q(y|x)$, tal que $\int q(y|x)dy = 1$. Esta función se puede interpretar al decir que cuando un proceso está en x , la f.d.p. genera un valor y de $q(y|x)$. Ahora, en el caso de que $q(y|x)$ satisfaga la condición de reversibilidad para todo (x, y) , la búsqueda terminaría (puesto que la reversibilidad es una condición suficiente para la convergencia de $f(\cdot)$); pero lo más probable es que no. Por ejemplo, se podría encontrarse con el caso

$$f(x)q(y|x) > f(y)q(x|y),$$

para algún (x, y) . En este caso, el proceso se mueve de x a y más a menudo que de y a x . Una forma de corregir esta condición es reducir los movimientos de x a y al introducir la probabilidad de movimiento

$$\alpha(x, y) < 1.$$

De esta manera, las transiciones de x a y se hacen de acuerdo a:

$$p_{MH}(x, y) := q(y|x)\alpha(x, y), \quad x \neq y, \quad (3.1.0.1)$$

donde $\alpha(x, y)$ no está todavía determinado.

Al calcular $\alpha(x, y)$ de tal forma que $p_{MH}(x, y)$ satisfaga la condición de reversibilidad, se tiene que:

$$f(x)p_{MH}(x, y) = f(y)p_{MH}(y, x), \quad (3.1.0.2)$$

$$= \pi(y)q(y, x)\alpha(y, x), \quad (3.1.0.3)$$

²Se utiliza la palabra “muestrear” en el siguiente sentido: una muestra de una distribución $P(x)$ es una sola realización de x , cuya distribución de probabilidad es $P(x)$ y no se refiere a una colección de realizaciones $\{x\}$.

³La irreducibilidad es la propiedad que, independientemente del estado actual, se puede alcanzar cualquier otro estado en tiempo finito.

⁴Se dice que una cadena de Márkov es aperiódica si todos sus estados son aperiódicos, es decir, sus periodos son iguales a 1. De lo contrario, se dice que la cadena es periódica.

⁵Una función $p(x, y)$ satisface la reversibilidad cuando $p(x)p(x, y) = p(y)p(y, x)$.

al definir $\alpha(y, x)$ tan grande como es posible, y puesto que es una probabilidad, su límite superior es 1, se obtiene que:

$$f(x)q(x, y)\alpha(x, y) = f(y)q(y, x).$$

Al despejar $\alpha(x, y)$, se obtiene:

$$\alpha(x, y) = \frac{f(y)q(y, x)}{f(x)q(x, y)}.$$

De esta manera, se ha demostrado que para poder garantizar la reversibilidad de $p_{MH}(x, y)$, la probabilidad de movimiento debe ser:

$$\alpha(x, y) = \begin{cases} \min \left[\frac{f(y)q(y, x)}{f(x)q(x, y)}, 1 \right] & \text{si } f(x)q(x, y) > 0, \\ 1 & \text{en otro caso,} \end{cases}$$

A continuación, se resume el algoritmo M-H tomado de Chib y Greenberg (1995) en el **Algoritmo 1**

Algoritmo 1 Algoritmo de Metropolis Hasting (MH)

Entrada: Dado una f.d.p. candidata $q(y|x)$ y un estado inicial $x^{(0)}$.

Salida: Muestras $f(\cdot)$, es decir, los valores $\{x^{(1)}, x^{(2)}, \dots, x^{(N)}\}$.

- 1: **Para** $j = 1, 2, \dots, N$.
- 2: **Generar** $y \sim q(\cdot|x^{(j)})$ y $u \sim \mathcal{U}(0, 1)$
- 3: **Si** $u \leq \alpha(x^{(j)}, y)$

$$x^{(j+1)} = y$$

donde:

$$\alpha(x^{(j)}, y) = \begin{cases} \min \left[\frac{f(y)q(y, x^{(j)})}{f(x^{(j)})q(x^{(j)}, y)}, 1 \right] & \text{si } f(x^{(j)})q(x^{(j)}, y) > 0, \\ 1 & \text{en otro caso,} \end{cases}$$

- 4: **Si no** $x^{(j+1)} = x^{(j)}$
-

Como en cualquier otro método de MCMC, sus valores de salida son vistos como obtenidos de la f.d.p. objetivo $\pi(\cdot)$; se consideran los datos después que la cadena ha pasado una etapa de transición o quemado, el cual es un periodo en el que las simulaciones obtenidas dependerán del valor tomado como estado inicial de las cadenas de Márkov, es decir, después de este periodo de transición, las simulaciones obtenidas no dependerán de los valores iniciales tomados.

3.2 Muestreo de Gibbs

Una vez más, se desea extraer muestras de un vector aleatorio $x = (x_1, \dots, x_n)$, con f.d.p. $f(x)$; sin embargo, resulta más fácil muestrear desde una f.d.p. condicional de X_i dado $X_j = x_j$ con $j \neq i$. Se denota a estas f.d.p. por:

$$f_i(x_i|x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n),$$

para $i = 1, \dots, n$. Si las simulaciones de f_i son fáciles de extraer, entonces se puede utilizar el muestreo de Gibbs para construir la cadena de Márkov X_1, X_2, \dots , con f.d.p. f .

El muestreo de Gibbs se reduce a un caso particular del algoritmo M-H visto en la sección anterior, al tomar como f.d.p. candidata generadora $q_{1 \rightarrow n}(y|x)$ por:

$$q_{1 \rightarrow n}(y|x) = f_1(y_1|x_2, \dots, x_n) f_2(y_2|y_1, x_3, \dots, x_n) \dots f_n(y_n|y_1, \dots, y_{n-1}),$$

y con probabilidad de aceptación de movimiento igual a uno.

De esta manera, Y_1 se obtiene utilizando la f.d.p. $f_1(y_1|x_2, \dots, x_n)$, Y_2 de $f_2(y_2|y_1, x_3, \dots, x_n)$ y así sucesivamente; y a diferencia del algoritmo de M-H, la simulación generada se acepta siempre, es decir, $X_{t+1} = Y$. El algoritmo de Gibbs es tomado de la página 218 de Kroese et al. (2014) y se resume en el **Algoritmo 2**.

Algoritmo 2 Algoritmo de Gibbs

Entrada: Dado las f.d.p. condicionales completas y un estado inicial $x^{(0)}$.

Salida: Muestras de $f(\cdot)$.

- 1: **Para** $t = 1, 2, \dots, n$.
 - 2: Dado el estado actual $X_t = x$, generar $Y = (Y_1, \dots, Y_n)$ como se expresa a continuación:
 - a. $Y_1 \sim f_1(y_1|x_2, \dots, x_n)$.
 - b. $Y_i \sim f_i(y_i|y_1, \dots, y_{i-1}, x_{i+1}, \dots, x_n)$. **para** $i = 2, \dots, n - 1$.
 - c. $Y_n \sim f_n(y_n|y_1, \dots, y_{n-1})$.
 - 3: **Actualizar** $X^{t+1} = Y$.
 - 4: **Devolver** los valores $\{Y_1, Y_2, \dots, Y_n\}$
-

A continuación, antes de describir el algoritmo de aceptación y rechazo de Metropolis Hasting, se introduce el muestreo de aceptación y rechazo.

3.3 Muestreo por aceptación y rechazo (AR)

En contraste a los métodos de MCMC, el método AR no genera una cadena de Márkov (generalmente independiente), a menos de que la correlación sea introducida artificialmente, como un mecanismo para reducir la varianza. El objetivo es generar muestras de una f.d.p. continua objetivo

$$\pi(x) = f(x)/K,$$

donde: $f(x)$ es una f.d.p. no-normalizada y K es una constante (posiblemente desconocida).

Sea $h(x)$ (conocida como función dominadora o de blanqueamiento) una f.d.p. que se puede obtener por algún método conocido y se supone que existe una constante c , tal que:

$$f(x) \leq ch(x)$$

para todo x . Para obtener las muestras de $\pi(\cdot)$, se sigue el **Algoritmo 3** tomado de Chib y Greenberg (1995).

Algoritmo 3 Algoritmo de aceptación y rechazo (AR)

Entrada: Dado la f.d.p. candidata $h(\cdot)$ y una constante c .

Salida: Muestras $\pi(\cdot)$.

1: Generar un candidato $z \sim h(\cdot)$ y un valor $u \sim \mathcal{U}(0, 1)$.

2: **Si** $u \leq f(z)/ch(z)$

Devolver z

3: **Si no** Volver a 1

A continuación, se muestra la segunda aplicación del algoritmo de M-H, una implementación del muestreo de Aceptación y Rechazo (AR) cuando la función dominadora no está disponible.

3.4 Algoritmo de aceptación y rechazo de Metropolis Hasting(ARMH)

El algoritmo es introducido en Tierney (1994), al introducir un paso de aceptación y rechazo (AR) para generar candidatos para un algoritmo M-H.

Una vez más, se desea obtener muestras de la f.d.p. $f(\theta|y) \propto f(y|\theta)\pi(\theta)$, y la f.p.d. $h(\theta|y)$ está disponible para el muestreo. En el método AR visto en la sección anterior, se tiene un requerimiento fundamental que exista una constante c , tal que la condición

$$D = \{\theta : f(y|\theta)\pi(\theta) \leq ch(\theta|y)\},$$

se mantenga para todo $\theta \in \Theta$; en cambio, el algoritmo ARMH es un procedimiento de muestreo MCMC en el cual la condición de dominación

$$f(y|\theta)\pi(\theta) \leq ch(\theta|y),$$

no es satisfecha para $\theta \in \Theta$; y por esta razón $h(\theta|y)$ se conoce como f.d.p. pseudo dominada, puesto que $ch(\theta|y)$ no domina la f.d.p. objetivo en D^c (por definición), se tiene que la función objetivo no se muestrea adecuadamente en esta región (D^c). Este problema se puede corregir aplicando un paso M-H a los candidatos que son generados, a través de un paso AR. El algoritmo ARMH se resume en el **Algoritmo 4**. (Tomado de Chib y Jeliazkov (2005))

Algoritmo 4 Algoritmo de aceptación rechazo de Metropolis Hasting (ARMH)

Entrada: Dado una f.d.p. pseudo dominada $h(\theta|y)$.

Salida: Muestras de $f(\theta|y)$.

1: **Paso AR:** Generar $\theta' \sim h(\theta|y)$; se acepta θ' con probabilidad

$$\alpha_{AR}(\theta' | y) = \min \left\{ 1, \frac{f(y|\theta')\pi(\theta')}{ch(\theta'|y)} \right\}$$

2: **Paso MH:** Dados el estado actual de la cadena θ y su valor propuesto θ' :

a. **Si** $\theta \in D$, se establece:

$$\alpha_{MH}(\theta, \theta' | y) = 1,$$

b. **Si** $\theta \in D^c$ y $\theta' \in D$, se establece:

$$\alpha_{MH}(\theta, \theta' | y) = \frac{ch(\theta|y)}{f(y|\theta)\pi(\theta)},$$

c. **Si** $\theta \in D^c$ y $\theta' \in D^c$, se establece:

$$\alpha_{MH}(\theta, \theta' | y) = \min \left\{ 1, \frac{f(y|\theta')\pi(\theta')h(\theta|y)}{f(y|\theta)\pi(\theta)h(\theta'|y)} \right\}.$$

3: **Devolver** θ' con probabilidad $\alpha(\theta, \theta' | y)$

CAPÍTULO 4

Estimación bayesiana del modelo SVM-TVP

En este capítulo se presenta un algoritmo basado en métodos MCMC introducido en Chan (2017) para estimar los parámetros del modelo SVM-TVP. Después, se propone una aplicación del trabajo de Chib (1995), basado en el muestreo de Gibbs para estimar la verosimilitud marginal del modelo SVM-TVP. Por último, se presentan los criterios de comparación entre modelos, utilizando: el factor bayesiano y los criterios de información de Akaike (AIC) y Schwarz (BIC).

Para la estimación del modelo SVM-TVP, existen varias razones para utilizar inferencia bayesiana y los métodos de MCMC. Primero, la función de verosimilitud es intratable, porque el modelo incluye ecuaciones de estado no lineales para la volatilidad estocástica, lo cual excluye el método de estimación de máxima verosimilitud. Además, no se tiene acceso a la forma analítica de la distribución a posteriori; segundo, al utilizar los métodos de MCMC, no solo se pueden muestrear simultáneamente el vector de parámetros $\psi = (\mu, \phi, \beta, \sigma^2, \Omega)'$; sino también las variables de estados $\gamma = (\gamma_1, \dots, \gamma_T)'$ y $h = (h_1, \dots, h_T)'$ se obtienen simultáneamente, y se pueden hacer inferencias sobre las variables de estados con la incertidumbre del vector de parámetros ψ .

Para el modelo SVM-TVP, al especificar la f.d.p. a priori $p(\psi)$ en (1.3.3.4), se puede obtener la f.d.p. a posteriori $p(\psi, \gamma, h|y)$. Existen varias formas de implementar el método de MCMC para explorar esta distribución a posteriori; sin embargo, se utiliza el muestreo de Gibbs, siguiendo la metodología expuesta por primera vez en Nakajima et al. (2011) y más tarde, en Chan (2017), a través del muestreo secuencial de las siguientes f.d.p. condicionales completas:

1. $p(h|y, \gamma, \mu, \phi, \beta, \sigma^2, \Omega) = p(h|y, \gamma, \mu, \phi, \beta, \sigma^2)$.
2. $p(\gamma|y, h, \mu, \phi, \beta, \sigma^2, \Omega) = p(\gamma|y, h, \Omega)$.
3. $p(\sigma^2, \Omega|y, \gamma, h, \mu, \phi, \beta) = p(\Omega|\gamma)p(\sigma^2|h, \mu, \phi, \beta)$.
4. $p(\mu, \phi|y, h, \gamma, \beta, \sigma^2, \Omega) = p(\mu, \phi|h, \sigma^2)$.
5. $p(\beta|y, h, \gamma, \mu, \phi, \sigma^2, \Omega) = p(\beta|y, h, \mu, \phi, \sigma^2)$.

Los detalles del procedimiento se ilustran, a continuación, en las siguientes secciones.

4.1 Muestreo de h

La f.d.p. condicional conjunta $p(h|y, \gamma, \mu, \phi, \sigma^2)$ es altamente dimensional y no lineal. Para obtener simulaciones desde esta f.d.p., se sigue el enfoque de Chan y Strachan (2014), en el cual se utiliza una aproximación gaussiana como f.d.p. propuesta (o candidata) en un algoritmo ARMH. Además, puesto que la matriz Hessiana de esta f.d.p. gaussiana propuesta es una matriz de banda, se puede explorar su estructura para calcular fácilmente su matriz inversa de covarianzas de los estados para obtener las simulaciones de los candidatos.

Primero, se discute una forma eficiente de obtener una aproximación gaussiana de

$$p(h|y, \gamma, \mu, \phi, \sigma^2).$$

Por el teorema de Bayes, se obtiene:

$$p(h|y, \gamma, \mu, \phi, \beta, \sigma^2) \propto p(y|h, \gamma)p(h|\mu, \phi, \beta, \sigma^2). \quad (4.1.0.1)$$

A continuación, se encuentran las f.d.p. del lado derecho de la expresión de arriba. Resulta que la f.d.p. a priori $p(h|\mu, \phi, \beta, \sigma^2)$ es gaussiana (la demostración se puede ver más abajo). Si se aproxima la función de verosimilitud $p(y|h, \gamma)$ por una f.d.p. gaussiana en h , inmediatamente, se obtiene una aproximación gaussiana de $p(h|y, \gamma, \mu, \phi, \beta, \sigma^2)$, para así, se pueda aproximar el logaritmo de la función de verosimilitud:

$$\log p(y|h, \gamma) = \sum_{t=1}^T \log p(y_t|h_t, \gamma_t)$$

en un punto $\tilde{h} = (\tilde{h}_1, \tilde{h}_2, \dots, \tilde{h}_T)' \in \mathbb{R}^T$, al utilizar una expansión de Taylor de segundo orden alrededor de \tilde{h} , para obtener:

$$\log p(y|h, \gamma) \approx \log p(y|\tilde{h}, \gamma) + (h - \tilde{h})' f - \frac{1}{2} (h - \tilde{h})' G (h - \tilde{h}), \quad (4.1.0.2)$$

$$\approx -\frac{1}{2} \left(h' G h - 2h'(f + G\tilde{h}) \right) + c_1, \quad (4.1.0.3)$$

con:

$$f_t = \frac{\partial}{\partial h_t} \log p(y_t|h_t, \gamma_t) \Big|_{h_t=\tilde{h}_t} \quad y \quad G_t = -\frac{\partial^2}{\partial h_t^2} \log p(y_t|h_t, \gamma_t) \Big|_{h_t=\tilde{h}_t},$$

donde: c_1 es una constante independiente de h , $f = (f_1, \dots, f_T)'$ y $G = \text{diagonal}(G_1, \dots, G_T)$ (donde $\text{diagonal}(\cdot)$ denota a la matriz diagonal).

En otras palabras, G es la negativa de la matriz Hessiana de la función de verosimilitud condicional en escala logarítmica evaluada en \tilde{h} . Además, se tiene que G es diagonal, puesto que el logaritmo de la f.d.p. condicional de y_t dado las variables latentes h_t y $\gamma_t = (\alpha_t, \beta_t)'$

está dado por:

$$\begin{aligned}\log p(y_t|h_t, \gamma_t) &= -\frac{1}{2} \log(2\pi) - \frac{1}{2} h_t - \frac{1}{2} e^{-h_t} ((y_t - \tau_t) - \alpha_t e^{h_t})^2, \\ &= -\frac{1}{2} \log(2\pi) - \frac{1}{2} h_t - \frac{1}{2} e^{-h_t} ((y_t - \tau_t)^2 - 2\alpha_t e^{h_t} (y_t - \tau_t) + \alpha_t^2 e^{2h_t}), \\ &= -\frac{1}{2} \log(2\pi) - \frac{1}{2} h_t - \frac{1}{2} (e^{-h_t} (y_t - \tau_t)^2 - 2\alpha_t (y_t - \tau_t) + \alpha_t^2 e^{h_t}),\end{aligned}$$

al derivar una y dos veces respecto a h_t , se obtiene:

$$\frac{\partial}{\partial h_t} \log p(y_t|h_t, \gamma_t) = -\frac{1}{2} - \frac{1}{2} \alpha_t^2 e^{h_t} + \frac{1}{2} e^{-h_t} (y_t - \tau_t)^2$$

y

$$\frac{\partial^2}{\partial h_t^2} \log p(y_t|h_t, \gamma_t) = -\frac{1}{2} \alpha_t^2 e^{h_t} - \frac{1}{2} e^{-h_t} (y_t - \tau_t)^2,$$

respectivamente.

A continuación, se encuentra la densidad a priori $p(h|\mu, \phi, \beta, \sigma^2)$. Para este fin, se establece:

$$H_\phi = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\phi & 1 & 0 & \cdots & 0 \\ 0 & -\phi & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -\phi & 1 \end{bmatrix}.$$

Se tiene que la matriz H_ϕ es inversible, puesto que su determinante es igual a 1, fácilmente verificable, ya que H_ϕ es bidiagonal. Ahora, al reescribir la ecuación de estados de h_t (ver ecuación 1.3.3.2) en forma matricial:

$$H_\phi h = \tilde{\delta}_h + \varepsilon^h, \quad \varepsilon^h \sim \mathcal{N}(0, S_h) \quad (4.1.0.4)$$

donde:

$$\tilde{\delta}_h = (\mu, (1 - \phi)\mu + \beta y_1, \dots, (1 - \phi)\mu + \beta y_{T-1})', \quad \varepsilon^h = (\varepsilon_1^h, \dots, \varepsilon_T^h)'$$

y

$$S_h = \text{diag} \left(\frac{\sigma^2}{(1 - \phi)}, \sigma^2, \dots, \sigma^2 \right),$$

puesto que la combinación lineal (afín) de una variable aleatoria normal sigue siendo normal¹, se obtiene:

$$h|\mu, \phi, \beta, \sigma^2 \sim \mathcal{N} \left(\delta_h, (H_\phi' S_h^{-1} H_\phi)^{-1} \right)$$

con densidad en escala logarítmica igual a:

$$\log p(h|\mu, \phi, \beta, \sigma^2) = -\frac{1}{2} \left(h^2 H_\phi' S_h^{-1} H_\phi h - 2h' H_\phi' S_h^{-1} H_\phi \delta_h \right) + c_2, \quad (4.1.0.5)$$

¹Se tiene que $AX + c \sim \mathcal{N}(A\mu + c, A\Sigma A')$, cuando $X \sim \mathcal{N}(\mu, \Sigma)$ (tomado del corolario 137 de Castro (2008)).

donde: $\delta_h = H_\phi^{-1}\tilde{\delta}_h$ y c_2 es una constante independiente de h . Finalmente, al combinar (4.1.0.3) y (4.1.0.5) en (4.1.0.1), se obtiene la siguiente expresión:

$$\log p(h|y, \gamma, \mu, \phi, \sigma^2) = \log p(y|h, \gamma) + \log p(h|\mu, \phi, \beta, \sigma^2) + c_3, \quad (4.1.0.6)$$

$$\propto -\frac{1}{2} \left(h' K_h h - 2h' k_h \right) + c_4, \quad (4.1.0.7)$$

donde: c_3 y c_4 son constantes independientes de h , con:

$$K_h = H'_\phi S_h^{-1} H_\phi + G$$

y

$$k_h = f + G\tilde{h} + H'_\phi S_h^{-1} H_\phi \delta_h.$$

Al utilizar los resultados de la página 238 de Kroese et al. (2014), se tiene que (4.1.0.7) es el logaritmo de la una f.d.p. normal:

$$h|y, \gamma, \mu, \phi, \beta, \sigma^2 \sim \mathcal{N} \left(\hat{h}, K_h^{-1} \right), \quad (4.1.0.8)$$

donde: $\hat{h} = K_h^{-1}k_h$. En otras palabras, $p(h|y, \gamma, \mu, \phi, \beta, \sigma^2)$ se puede aproximar por una densidad gaussiana con media \hat{h} y matriz de precisión K_h . Esta aproximación gaussiana se puede utilizar como f.d.p. candidata en el algoritmo ARMH. Ahora, solamente faltaría por seleccionar el punto \tilde{h} para la expansión de Taylor en (4.1.0.3).

Para finalizar, se puede estimar \tilde{h} al maximizar $p(h|y, \gamma, \mu, \phi, \beta, \sigma^2)$, o lo que es equivalente a maximizar al $\log p(h|y, \gamma, \mu, \phi, \beta, \sigma^2)$, al utilizar el método de Newton-Raphson (ver Kroese et al. (2013), páginas 688-689).

Se obtiene de (4.1.0.7) la matriz Hessiana negativa de $\log p(h|y, \gamma, \mu, \phi, \beta, \sigma^2)$ evaluada en $h = \tilde{h}$ es K_h y el gradiente evaluado en $h = \tilde{h}$ es $-K_h\tilde{h} + k_h$. De esta manera, se puede complementar el método de Newton-Raphson, como se expresa a continuación: se inicializa con $h = h^{(1)}$, para algún vector constante $h^{(1)}$. Al reemplazar $\tilde{h} = h^{(t)}$ en la evaluación de K_h y k_h , se obtiene:

$$\begin{aligned} h^{(t+1)} &= h^{(t)} + K_h^{-1}(K_h h^{(t)} + k_h), \\ &= K_h^{-1}k_h, \end{aligned}$$

para $t = 1, 2, \dots$. Se repite este procedimiento hasta que un criterio de convergencia sea alcanzado, por ejemplo, cuando:

$$\|h^{(t+1)} - h^{(t)}\| < \epsilon,$$

para algún nivel de tolerancia prefijado ϵ .

Ahora, antes de abordar el muestreo de γ , se enuncia el siguiente resultado de las f.d.p. condicionales a posteriori para modelos lineales (tomado de Kroese et al. (2014), páginas 239 y 240).

Teorema 1 (Corolario 8.1 de Kroese et al. (2014)). *Se considera el modelo bayesiano:*

$$\beta \sim \mathcal{N}(\beta_0, \Sigma_0), \quad (4.1.0.9)$$

$$\sigma^2 \sim \mathcal{IG}(\alpha_0, \lambda_0), \quad (4.1.0.10)$$

$$(y - X\beta | \beta, \sigma^2) \sim \mathcal{N}(0, \sigma^2 R), \quad (4.1.0.11)$$

donde: $y - X\beta$ es una transformación afín de y , R y Σ_0 son conocidas, β_0 es un vector fijo, α_0 y λ_0 son constantes fijas. Entonces,

$$(\beta | \sigma^2, y) \sim \mathcal{N}(\mu, D),$$

donde: $\mu = D (X' R' y / \sigma^2 + \Sigma_0^{-1} \beta_0)$, con: $D = (X' R^{-1} X / \sigma^2 + \Sigma_0^{-1})^{-1}$ y

$$(\sigma^2 | \beta, y) \sim \mathcal{IG}(\alpha_0 + n/2, (y - X\beta)' R' (y - X\beta) / 2 + \lambda_0).$$

4.2 Muestreo de γ

A continuación, se aborda una forma eficiente de obtener muestras de $p(\gamma | y, h, \Omega)$. Al reescribir (1.3.3.1), como se expresa:

$$y_t = z_t' \gamma_t + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, e^{h_t}), \quad (4.2.0.12)$$

donde: $z_t = (e^{h_t}, 1)'$ y $\gamma_t = (\alpha_t, \beta_t)'$. Por lo tanto, (4.2.0.12) y (1.3.3.3) definen un espacio de estados lineal gaussiano con respecto a γ_t , como estado. Por consiguiente, se puede obtener simulaciones de $p(\gamma | y, h, \Omega)$ por medio del filtro de Kalman ² o por un procedimiento más eficiente basado en inferencia bayesiana tomado del Apéndice A de Chan (2017). Se elige el segundo enfoque, para este fin, se reescribe (4.2.0.12) en forma matricial, como se expresa a continuación:

$$y = Z\gamma + \varepsilon^y, \quad \varepsilon^y \sim \mathcal{N}(0, S_y), \quad (4.2.0.13)$$

donde: $\varepsilon^y = (\varepsilon_1^y, \dots, \varepsilon_T^y)$,

$$S_y = \begin{bmatrix} e^{h_1} & 0 & \dots & 0 \\ 0 & e^{h_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{h_T} \end{bmatrix}, \quad Z = \begin{bmatrix} z_1' & 0 & \dots & 0 \\ 0 & z_2' & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & z_T' \end{bmatrix}.$$

Por otro lado, para encontrar la densidad a priori de γ , se reescribe la ecuación de estados de γ_t (ver la ecuación (1.3.3.3)) en forma matricial:

$$H\gamma = \tilde{\delta}_\gamma + \varepsilon^y, \quad \varepsilon^y \sim \mathcal{N}(0, S_\gamma), \quad (4.2.0.14)$$

²Se explica de manera detallada el filtro de Kalman en la página 326 de Shumway y Stoffer (2011).

donde: $\tilde{\delta}_\gamma = (\gamma'_0, 0, \dots, 0)'$, $\varepsilon^y = ((\varepsilon_1^\gamma)', \dots, (\varepsilon_T^\gamma)')$, $S_\gamma = \text{diag}(\Omega_0, \Omega, \dots, \Omega)$ y H es la matriz de diferencias:

$$H = \begin{bmatrix} I_{k+1} & 0 & 0 & \cdots & 0 \\ -I_{k+1} & I_{k+1} & 0 & \cdots & 0 \\ 0 & -I_{k+1} & I_{k+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -I_{k+1} & I_{k+1} \end{bmatrix}.$$

Al aplicar el corolario 137 de Castro (2008) (Combinaciones lineales de una v.a. normal) a (4.2.0.14), se obtiene:

$$(\gamma|\Omega) \sim \mathcal{N}\left(\delta_\gamma, (H' S_\gamma^{-1} H)^{-1}\right), \quad (4.2.0.15)$$

donde: $\delta_\gamma = H^{-1} \tilde{\delta}_\gamma$. Ahora, al utilizar el Corolario 8.1 de Kroese et al. (2014) (Teorema 1), enunciado antes de empezar esta sección con (4.2.0.14) y (4.2.0.15), se obtiene:

$$(\gamma|y, h, \Omega) \sim \mathcal{N}(\hat{\gamma}, D_\gamma), \quad (4.2.0.16)$$

donde: $D_\gamma^{-1} = H' S_\gamma^{-1} H + Z' S_y^{-1} Z$ y $\hat{\gamma} = D_\gamma (H' S_\gamma^{-1} H \delta_\gamma + Z' S_y^{-1} y)$.

4.3 Muestreo de Ω

El siguiente resultado se ha tomado de la página 114 de Nakajima et al. (2011), para obtener la densidad condicional a posteriori de Ω . Se asume que la f.d.p. a priori de Ω es

$$\Omega \sim \mathcal{IW}(v_\Omega, S_\Omega),$$

donde \mathcal{IW} es la distribución inversa de Wishart. A continuación, se obtiene la distribución condicional a posteriori de Ω , que se puede deducir, como se explica a continuación:

$$p(\Omega|\gamma) \propto p(\gamma|\Omega)p(\Omega),$$

$$\begin{aligned} &\propto |\Omega|^{-\frac{v_\Omega+p+1}{2}} \exp\left\{-\frac{1}{2} \text{Tr}(S_\Omega^{-1} \Omega^{-1})\right\} \prod_{t=2}^T \frac{1}{|\Omega|^{1/2}} \exp\left\{-\frac{1}{2} (\gamma_t - \gamma_{t-1})' \Omega^{-1} (\gamma_t - \gamma_{t-1})\right\}, \\ &\propto |\Omega|^{-\frac{(\tilde{v}_\Omega+p+1)}{2}} \exp\left\{-\frac{1}{2} \text{Tr}(\tilde{S}_\Omega \Omega^{-1})\right\}, \end{aligned} \quad (4.3.0.17)$$

donde:

$$\tilde{v}_\Omega = v_\Omega + T - 1 \quad \text{y} \quad \tilde{S}_\Omega = S_\Omega + \sum_{t=2}^T (\gamma_t - \gamma_{t-1})(\gamma_t - \gamma_{t-1})'.$$

Nótese que la f.d.p. a posteriori para Ω depende, solamente, de γ y de (4.3.0.17); se concluye que la f.d.p. a posteriori de Ω sigue una distribución inversa de Wishart, tal como, se expresa a continuación:

$$\Omega|\gamma \sim \mathcal{IW}(v_\Omega + T - 1, \tilde{S}_\Omega). \quad (4.3.0.18)$$

4.4 Muestreo de σ^2

El siguiente resultado se ha tomado de la página 116 de Nakajima et al. (2011), para obtener la densidad condicional a posteriori de σ^2 . Se asume que f.d.p. a priori de σ^2 es

$$\sigma^2 \sim \mathcal{IG}(v_{\sigma^2}, S_{\sigma^2}), \quad (4.4.0.19)$$

donde \mathcal{IG} se refiere a la distribución inversa de Gamma. Al reescribir (1.3.3.2) en forma matricial:

$$Z_{\beta} = X_{\beta}\beta + \varepsilon^h, \quad \varepsilon^h \sim \mathcal{N}(0, \sigma^2 I),$$

donde: $Z_{\beta} = \left(\sqrt{1 - \phi^2}(h_1 - \mu), h_2 - \phi h_1 - \mu(1 - \phi), \dots, h_T - \phi h_{T-1} - \mu(1 - \phi) \right)'$ y $X_{\beta} = (0, y_1, \dots, y_{T-1})'$, por consiguiente:

$$Z_{\beta} - X_{\beta}\beta | y, h, \mu, \phi, \beta, \sigma^2 \sim \mathcal{N}(0, \sigma^2 I). \quad (4.4.0.20)$$

Una vez más, al utilizar el Corolario 8.1. de Kroese et al. (2014) con (4.4.0.19) y (4.4.0.20), se obtiene la siguiente expresión:

$$\sigma^2 | h, \mu, \phi \sim \mathcal{IG}\left(v_{\sigma^2} + \frac{T}{2}, \tilde{S}_{\sigma^2}\right), \quad (4.4.0.21)$$

donde: $\tilde{S}_{\sigma^2} = S_{\sigma^2} + \left[(1 - \phi^2)(h_1 - \mu)^2 + \sum_{t=2}^T (h_t - \mu - \phi(h_{t-1} - \mu) - \beta y_{t-1})^2 / 2 \right]$.

4.5 Muestreo de β

El siguiente resultado se ha tomado de la página 10 de Chan (2017), para obtener la densidad condicional a posteriori de σ^2 . Se asume que la f.d.p. a priori de β es

$$\beta \sim \mathcal{N}(0, V_0). \quad (4.5.0.22)$$

Al reescribir (1.3.3.2) en forma matricial, como se expresa a continuación:

$$Z_{\beta} = X_{\beta}\beta + \varepsilon^h, \quad \varepsilon^h \sim \mathcal{N}(0, \sigma^2 I),$$

donde: $Z_{\beta} = (h_2 - \phi h_1 - \mu(1 - \phi), \dots, h_T - \phi h_{T-1} - \mu(1 - \phi))'$ y $X_{\beta} = (y_1, \dots, y_{T-1})'$; por consiguiente, se obtiene la siguiente expresión:

$$Z_{\beta} - X_{\beta}\beta | y, h, \mu, \phi, \sigma^2 \sim \mathcal{N}(0, \sigma^2 I). \quad (4.5.0.23)$$

Al utilizar el Corolario 8.1 de Kroese et al. (2014) (Teorema 1) con (4.5.0.22) y (4.5.0.23), se obtiene:

$$\beta | y, h, \mu, \phi, \sigma^2 \sim \mathcal{N}\left(\hat{\beta}, D_{\beta}\right), \quad (4.5.0.24)$$

donde: $D_{\beta}^{-1} = V_{\beta}^{-1} + X_{\beta}'X_{\beta}/\sigma^2$ y $\hat{\beta} = D_{\beta}(V_{\beta}^{-1}\beta_0 + X_{\beta}'Z_{\beta}/\sigma^2)$.

4.6 Muestreo de μ y ϕ

A continuación, se muestra el procedimiento utilizado en Chan (2017), para obtener la densidad condicional a posteriori de μ y ϕ . Se supone que las f.d.p. a priori para μ y ϕ son

$$\mu \sim \mathcal{N}(\mu_0, V_\mu)$$

y

$$\phi \sim \mathcal{N}(\phi_0, V_\phi) \mathbb{1}(|\phi| < 1),$$

respectivamente. Se implementa el método de cadenas de Márkov de Metropolis Hasting, para extraer las simulaciones de la siguiente f.d.p. a posteriori de μ y ϕ en escala logarítmica:

$$\begin{aligned} \log p(\mu, \phi | h, \beta, \sigma^2) &= \log(p(h | \mu, \phi, \beta, \sigma^2) p(\mu, \phi | \beta, \sigma^2)), \\ &= \log(p(h | \mu, \phi, \beta, \sigma^2) p(\mu) p(\phi)), \\ &= -\frac{(1-\phi)^2}{2\sigma^2} (h_1 - \mu)^2 - \frac{1}{2\sigma^2} \sum_{t=2}^T (h_t - \phi h_{t-1} - \mu(1-\phi) - \beta y_{t-1})^2 \\ &\quad - \frac{T}{2} \log(2\pi\sigma^2) - \frac{1}{2} \log(2\pi V_{\phi_0}) - \frac{1}{2V_{\phi_0}} (\phi - \phi_0)^2 - \frac{1}{2} \log(2\pi V_{\mu_0}) \\ &\quad - \frac{1}{2V_{\mu_0}} (\mu - \mu_0)^2, \end{aligned} \tag{4.6.0.25}$$

puesto que la densidad a posteriori no se puede expresar de una manera analíticamente simple.

Para obtener muestras de la f.d.p. a posteriori (4.6.0.25), se utiliza como densidad propuesta ($q(\cdot | \cdot)$) a una t de Student bivalente $t_v(u, \Sigma)$, con media u igual a la moda de la f.d.p. a posteriori $p(\mu, \phi | h, y, \beta, \sigma^2)$:

$$u = \underset{\mu, \phi}{\operatorname{argm\acute{a}x}} \log p(\mu, \phi | h, y, \beta, \sigma^2),$$

(es decir, que el método de estimación de máxima a posteriori estima a los parámetros como la moda de la distribución a posteriori de esta v.a.), la cual se puede calcular por medio del método de Newton-Raphson (con BHHH). La varianza (Σ) es la inversa de la Hessiana de (4.6.0.25) con signo negativo evaluado en u . Este enfoque para especificar la densidad candidata o generadora fue introducido en Chib y Jeliazkov (2001), páginas 274 y 275.

A continuación, se enuncia el teorema de la transformación tomado de Kroese et al. (2014) (página 85, Teorema 3.5.), el cual será utilizado para extraer las simulaciones conjuntas de μ y $|\phi| < 1$.

Teorema 2 (Teorema de la transformación). *Sea X un vector aleatorio de dimensión n con f.d.p. f_X y g una función de \mathbb{R}^n a \mathbb{R}^n con inversa g^{-1} . Entonces, $Z = g(X)$ tiene como f.d.p. a:*

$$f_Z(z) = f_X(g^{-1}(z)) |J_{g^{-1}}(z)|, \quad z \in \mathbb{R}_n,$$

donde: $|J_{g^{-1}}(z)| = 1/|J_g(x)|$ y $|J_g(z)|$ denota a la determinante de la matriz Jacobiana evaluada en z .

4.6.1 Restricción del parámetro ϕ

En el modelo SVM-TVP se tiene la siguiente suposición en (1.3.3.2):

$$|\phi| < 1.$$

Es posible utilizar la transformación:

$$\eta = \operatorname{arctanh}(\phi), \quad -\infty < \eta < \infty, \quad (4.6.1.1)$$

donde el parámetro restringido ϕ se puede obtener de vuelta con:

$$\phi = \tanh(\eta), \quad -1 \leq \phi \leq 1. \quad (4.6.1.2)$$

Con estas transformaciones, las ϕ propuestas se generan en dos pasos: primero, se genera μ y η de la f.d.p. generadora ($t_v(u, \Sigma)$); en el segundo paso, se obtiene de regreso a ϕ utilizando (4.6.1.2); y si se requiere continuar con el algoritmo de M-H, se regresa a η utilizando (4.6.1.1).

Además, se debe conocer la f.d.p. generadora conjunta de μ y ϕ para calcular la probabilidad de aceptación al emplear el algoritmo M-H. Al conocer la f.d.p. generadora conjunta de μ y η , se puede encontrar la f.d.p. generadora conjunta de μ y ϕ , utilizando la transformación:

$$g : (\mu, \eta) \longrightarrow (\mu, \phi = \tanh(\eta)).$$

Entonces, la matriz Jacobiana de $g^{-1} : (\mu, \phi) \longrightarrow (\mu, \operatorname{arctanh}(\eta))$ es

$$J_{g^{-1}}(\mu, \phi) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{1-\phi^2} \end{pmatrix}$$

con determinante $|J_{g^{-1}}(\mu, \phi)| = 1/(1-\phi^2)$. Al utilizar el teorema de la transformación, se concluye que μ y ϕ tienen como f.d.p. generadora conjunta a

$$t_v(u, \Sigma) / (1 - \phi^2),$$

donde: $t_v(u, \Sigma)$ es la f.d.p. conjunta de μ y η .

4.6.2 Cálculo de los parámetros de la función propuesta generadora

A continuación, en esta subsección se mostrarán los cálculos, para encontrar los parámetros u , Σ y v , de la f.d.p. propuesta generadora; primero, se enuncia rápidamente el algoritmo de Newton-Raphson con BHHH para encontrar u ; segundo, se brinda una aproximación del parámetro Σ ; y finalmente, se establecen los grados de libertad v , al compararlos con su tasa de aceptación en el algoritmo M-H.

El algoritmo de Berndt-Hall-Hall-Hausman (BHHH) fue introducido en Berndt et al. (1974), el cual es un algoritmo de optimización numérica similar al algoritmo de

Newton-Raphson; y se lo utiliza para encontrar el máximo del logaritmo de la función de verosimilitud (Q), donde:

$$Q = \sum_{t=1}^T Q_t.$$

En particular, el procedimiento iterativo se define, como se expresa a continuación:

$$u_{k+1} = u_k + \lambda_k (-H_k^{-1} g_k), \quad (4.6.2.1)$$

con:

$$-H_k = \sum_{t=1}^T g_t g_t' \quad (4.6.2.2)$$

y

$$g_k = \sum_{t=1}^T g_t, \quad (4.6.2.3)$$

donde: λ_k es un parámetro (llamado tamaño de paso) y g_t es el gradiente de Q_t . De esta manera, el algoritmo de BHHH se diferencia del Newton-Raphson, al aproximar la Hessiana negativa con el producto del gradiente por su transpuesta.

En particular, la función de verosimilitud está dada por (4.6.0.25), se la puede reescribir como una sumatoria de Q_t , utilizando:

$$Q_1 = -\frac{(1-\phi)^2}{2\sigma^2} (h_1 - \mu)^2 - \frac{1}{2V_{\phi_0}} (\phi - \phi_0)^2 - \frac{1}{2V_{\mu_0}} (\mu - \mu_0)^2 \quad (4.6.2.4)$$

y

$$Q_t = -\frac{1}{2\sigma_2} (h_t - \phi h_{t-1} - \mu(1-\phi) - \beta y_{t-1})^2, \quad (4.6.2.5)$$

para $2 \leq t \leq T$. Al realizar el cambio de variable, dado por: (4.6.1.1) en (4.6.2.4) y (4.6.2.5), se obtiene las siguientes expresiones:

$$Q_1 = -\frac{\text{sech}^2 \eta}{2\sigma^2} (h_1 - \mu)^2 - \frac{1}{2V_{\phi_0}} (\tanh \eta - \phi_0)^2 - \frac{1}{2V_{\mu_0}} (\mu - \mu_0)^2 \quad (4.6.2.6)$$

y

$$Q_t = -\frac{1}{2\sigma^2} (h_t - h_{t-1} \tanh \eta - \mu(1 - \tanh \eta) - \beta y_{t-1})^2, \quad (4.6.2.7)$$

para $2 \leq t \leq T$. Se obtiene el gradiente de Q al derivar (4.6.2.6) y (4.6.2.7), la cual está dada por:

$$g_1 = \begin{pmatrix} \frac{\partial Q_1}{\partial \mu} \\ \frac{\partial Q_1}{\partial \eta} \end{pmatrix} = \begin{pmatrix} \frac{\text{sech}^2 \eta}{\sigma^2} (h_1 - \mu) - \frac{(\mu - \mu_0)}{V_{\mu_0}} \\ \text{sech}^2 \eta \tanh \eta \left(\frac{(h_1 - \mu)^2}{\sigma^2} - \frac{1}{V_{\phi_0}} \right) + \frac{\phi_0}{V_{\phi_0}} \text{sech}^2 \eta \end{pmatrix} \quad (4.6.2.8)$$

y

$$g_t = \begin{pmatrix} \frac{\partial Q_t}{\partial \mu} \\ \frac{\partial Q_t}{\partial \eta} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma^2} (h_t - h_{t-1} \tanh \eta - \mu(1 - \tanh \eta) - \beta y_{t-1}) \\ \frac{\text{sech}^2 \eta (h_{t-1} - \mu)}{\sigma^2} (h_t - \mu - \tanh \eta (h_{t-1} - \mu) - \beta y_{t-1}) \end{pmatrix}, \quad (4.6.2.9)$$

para $2 \leq t \leq T$. De esta manera, se puede calcular iterativamente u , al reemplazar (4.6.2.8) y (4.6.2.9) en (4.6.2.2), (4.6.2.3) y (4.6.2.1). Además, se puede aproximar la Hessiana de (4.6.0.25), utilizando (4.6.2.2).

Finalmente, solo faltaría establecer v (grados de libertad). En el Cuadro 4.1 se muestran las tasas de aceptación para cada grado de libertad, en la cual se puede observar que cuando aumentan los grados de libertad, aumenta también la tasa de aceptación de las simulaciones generadas por la función candidata generadora. Especialmente, a partir de 5 grados de libertad; se observa que la tasa de aceptación, solamente, aumentó 1,12 % cuando los grados de libertad varían de 5 a 10, debido a que a medida que crece el número de grados de libertad, la distribución t se aproxima a la distribución normal (a v también se la conoce como el parámetro de normalidad). Se decide tomar $v = 5$, aunque se podría tomar cualquier valor mayor a 5; pero no se notaría una diferencia significativa en los resultados del muestreo de μ y ϕ .

Cuadro 4.1: Tasa de aceptación de la función propuesta generadora al extraer las simulaciones de μ y ϕ .

v	Tasa de aceptación
1	51,73 %
2	59,69 %
3	62,74 %
4	63,55 %
5	65,55 %
7	66,29 %
10	66,67 %

4.7 Verosimilitud marginal

En la presente sección se utiliza la estimación bayesiana por medio del muestreo de Gibbs, para desarrollar un enfoque simple que permita calcular la densidad marginal de los datos (verosimilitud marginal), a partir de simulaciones obtenidas de las f.d.p. a posteriori de los parámetros. Consecuentemente, se pueden calcular los factores bayesianos para realizar las respectivas comparaciones entre modelos; sin embargo, estos cálculos son extremadamente difíciles. El acercamiento utilizado aquí está basado en el trabajo expuesto en Chib (1995), el cual explota el hecho de que la densidad marginal se puede expresar como el número de veces de la f.d.p. a priori de la función de verosimilitud sobre la f.d.p. a posteriori. Se demuestra que una estimación de una f.d.p. a posteriori está disponible si todas las simulaciones de las condicionales completas lo están. Para mejorar la precisión de la densidad a posteriori, esta se estima en un punto de alta densidad. A continuación, se aplican estas ideas al modelo SVM-TVP.

4.7.1 Verosimilitud marginal del modelo SVM-TVP

En Newton y Raftery (1994) se muestra que la función de verosimilitud marginal (equivalente a la densidad marginal de y) bajo el modelo M_k , está dado por:

$$m(y|M_k) = \int p(y|\psi, M_k)p(\psi|M_k)d\psi,$$

donde: $p(y|\psi, M_k)$ es la función de densidad de los datos $y = (y_1, \dots, y_T)$ bajo el modelo M_k , para $k = 1, 2, \dots, K$, dado el vector de parámetros especificados del modelo ψ ; $p(\psi|M_k)$ es la densidad a priori de ψ ; y $p(\psi|y, M_k)$ es la f.d.p. a posteriori.

Al suprimir el índice k para el modelo, se considera a $p(y|\psi)$ como la función de verosimilitud para un modelo dado y $p(\psi)$ su densidad a priori. Se supone que para un conjunto de bloques de parámetros $\psi = ((\mu, \phi), \beta, (\sigma^2, \Omega))$ se conocen las siguientes densidades condicionales completas:

$$p(\mu, \phi|y, h, \gamma, \beta, \sigma^2, \Omega), \quad (4.7.1.1)$$

$$p(\beta|y, h, \mu, \phi, \sigma^2, \Omega), \quad (4.7.1.2)$$

$$p(\sigma^2, \Omega|y, h, \gamma, \mu, \phi, \beta), \quad (4.7.1.3)$$

$$p(h|y, \gamma, \mu, \phi, \beta, \sigma^2, \Omega) \text{ y} \quad (4.7.1.4)$$

$$p(\gamma|y, h, \mu, \phi, \beta, \sigma^2, \Omega); \quad (4.7.1.5)$$

para aplicarlas al muestreo de Gibbs.

El objetivo es calcular la densidad marginal $m(y|M_k)$, al utilizar las simulaciones

$$\{(\mu^{(g)}, \phi^{(g)}), \beta^{(g)}, (\sigma^{2(g)}, \Omega^{(g)}), h^{(g)}, \gamma^{(g)}\}$$

generadas por (4.7.1.1), (4.7.1.2), (4.7.1.3), (4.7.1.4) y (4.7.1.5), respectivamente.

El enfoque utilizado aquí consiste en dos ideas. La primera, utilizando el teorema de Bayes, $m(y)$ se puede expresar como:

$$m(y) = \frac{p(y|\psi)p(\psi)}{p(\psi|y)}. \quad (4.7.1.6)$$

A esta simple expresión, se la conoce como la identidad básica de la función de verosimilitud marginal (BML). La segunda, para una ψ dada (ψ^*), la función de verosimilitud y la f.d.p. a posteriori se estiman en ψ^* , y se denotan por $p(y|\psi^*)$ y $p(\psi^*|y)$, respectivamente, las cuales se pueden estimar utilizando las densidades completas ((4.7.1.1), (4.7.1.2), (4.7.1.3), (4.7.1.4) y (4.7.1.5)). La técnica para hacer esto se discutirá más adelante. Por otro lado, si la función de verosimilitud y la f.d.p. a posteriori se estiman en ψ^* , estas se denotan por: $\hat{p}(y|\psi^*)$ y $\hat{p}(\psi^*|y)$, respectivamente, entonces la estimación propuesta de la densidad marginal sobre una escala logarítmica es

$$\log \hat{m}(y) = \log \hat{p}(y|\psi^*) + \log p(\psi^*) - \log \hat{p}(\psi^*|y). \quad (4.7.1.7)$$

4.7.1.1 Elección de ψ^*

La expresión BML en (4.7.1.6) se cumple para cualquier ψ , por esta razón la elección del punto no es crítica; pero las consideraciones de eficiencia dictaminan que para un número dado de muestras de la f.d.p. a posteriori, la densidad se estima con mayor precisión en un punto de alta densidad, donde más muestras estén disponibles que en un punto en las colas. A continuación, se muestra como la función de verosimilitud y la densidad a posteriori se pueden estimar utilizando el muestreo de Gibbs.

4.7.1.2 Estimación de la verosimilitud $p(y|\psi^*)$

La función de verosimilitud $p(y|\psi^*)$ (también llamada verosimilitud integral) se puede obtener al utilizar la regla de producto de probabilidades:

$$p(y|\psi) = p(y_1|\psi) \prod_{t=1}^{T-1} p(y_{t+1}|Y_{1:t}, \psi),$$

donde: $Y_{1:t} = \{y_1, \dots, y_t\}$ y $p(y_{t+1}|Y_{1:t}, \psi)$ es la densidad predictiva, la cual es básicamente la densidad de y_{t+1} a un paso. Por definición, la densidad predictiva se define por:

$$p(y_{t+1}|Y_t, \psi) = \int p(y_{t+1}|Y_t, h_{t+1}, \gamma_{t+1}, \psi) p(h_{t+1}, \gamma_{t+1}|Y_t, \gamma_t, h_t, \psi) p(h_t, \gamma_t|Y_t, \psi) dh_{t+1} d\gamma_{t+1} dh_t d\gamma_t.$$

Entonces, la verosimilitud integral se puede estimar utilizando las densidades predictivas a un paso, como se expresa a continuación:

$$\log \widehat{p}(y|\psi^*) = \frac{1}{M} \sum_{i=1}^M p(y_t|h_t^{(i)}, \gamma_t^{(i)}, \psi^*), \quad (4.7.1.8)$$

donde: las simulaciones de la f.d.p. $h_t^{(i)}, \gamma_t^{(i)} | \gamma_{t-1}^{(i)}, h_{t-1}^{(i)}, \psi^*$, se pueden obtener utilizando las ecuaciones de estados (1.3.3.2) y (1.3.3.3). Por último, $h_{t-1}^{(i)}, \gamma_{t-1}^{(i)} | Y_{t-1}, \psi^*$ se puede obtener por el muestreo de Gibbs utilizando (4.1.0.8) y (4.2.0.16), con ψ^* fijo. Se observa que en la ecuación (4.7.1.8) se utiliza $p(y_t|h_t^{(i)}, \gamma_t^{(i)}, \psi^*)$ en vez de $p(y_t|Y_{t-1}, h_t^{(i)}, \gamma_t^{(i)}, \psi^*)$; esto se debe a que y_t en la ecuación de observación (1.3.3.1) solo depende de los estados y de los parámetros, que varían en el tiempo.

4.7.1.3 Estimación de $p(\psi^*|y)$

Ahora, el objetivo es estimar $p(\psi^*|y)$, el cual se puede expresar por:

$$p(\psi^*|y) = p(\mu^*, \phi^*|y) p(\sigma^{2*}, \Omega^*|y, \mu^*, \phi^*) p(\beta^*|y, \mu^*, \phi^*, \sigma^{2*}, \Omega^*), \quad (4.7.1.9)$$

donde:

$$p(\mu^*, \phi^* | y) = \int p(\mu^*, \phi^* | y, h, \gamma, \beta, \sigma^2, \Omega) p(h, \gamma, \beta, \sigma^2, \Omega | y) d\beta d\sigma^2 d\Omega dh d\gamma, \quad (4.7.1.10)$$

$$p(\sigma^{2*}, \Omega^* | y, \mu^*, \phi^*) = \int p(\sigma^{2*}, \Omega^* | y, h, \gamma, \mu^*, \phi^*, \beta) p(h, \gamma, \beta | y, \mu^*, \phi^*) d\beta dh d\gamma \quad (4.7.1.11)$$

y

$$p(\beta^* | y, \mu^*, \phi^*, \sigma^{2*}, \Omega^*) = \int p(\beta^* | y, h, \gamma, \mu^*, \phi^*, \sigma^{2*}, \Omega^*) p(h, \gamma | y, \mu^*, \phi^*, \sigma^{2*}, \Omega^*) dh d\gamma \quad (4.7.1.12)$$

son las densidades condicionales reducidas.

Entonces, se tiene que una estimación apropiada de (4.7.1.10) es el promedio de la f.d.p. condicional completa de (μ, ϕ) evaluada en (μ^*, ϕ^*) , como se expresa a continuación:

$$\hat{p}(\mu^*, \phi^* | y) = \frac{1}{G} \sum_{g=1}^G p(\mu^*, \phi^* | y, h^{(g)}, \gamma^{(g)}, \beta^{(g)}, \sigma^{2(g)}, \Omega^{(g)}), \quad (4.7.1.13)$$

donde: las simulaciones $\{h^{(g)}, \gamma^{(g)}, \beta^{(g)}, (\sigma^{2(g)}, \Omega^{(g)})\}$ se obtienen al aplicar el muestreo de Gibbs con:

$$\begin{aligned} p(h|y, \gamma, \mu, \phi, \beta, \sigma^2, \Omega) &= p(h|y, \gamma, \mu, \phi, \beta, \sigma^2), \\ p(\gamma|y, h, \beta, \sigma^2, \Omega) &= p(\gamma|y, h, \Omega), \\ p(\beta|y, h, \gamma, \mu, \phi, \sigma^2, \Omega) &= p(\beta|y, h, \mu, \phi, \sigma^2), \\ p(\mu, \phi|y, h, \gamma, \beta, \sigma^2, \Omega) &= p(\mu, \phi|h, \sigma^2) \text{ y} \\ p(\sigma^2, \Omega|y, h, \gamma, \mu, \phi, \beta) &= p(\Omega|\gamma) p(\sigma^2|h, \mu, \phi). \end{aligned} \quad (4.7.1.14)$$

La estimación en (4.7.1.10) es consistente, es decir, $\hat{p}(\mu^*, \phi^* | y) \rightarrow p(\mu^*, \phi^* | y)$ cuando $G \rightarrow \infty$, casi seguramente, como consecuencia de que la ley de los grandes números, se cumple para cualquier cadena ergódica (Teorema 3 de Tierney (1994)).

Una técnica similar, se puede utilizar para obtener la f.d.p. condicional reducida de $p(\sigma^{2*}, \Omega^* | y, \mu^*, \phi^*)$ en (4.7.1.11); pero si se emplea un muestreo de Gibbs utilizando las distribuciones marginales completas de $h, \gamma, \Omega, \sigma^2, \beta, \mu$ y ϕ (ver las ecuaciones (4.1.0.8), (4.2.0.16), (4.3.0.18), (4.4.0.21), y (4.5.0.24) y la sección 4.6, respectivamente), se obtienen las simulaciones de la f.d.p. $h, \gamma, \beta | y$ y no de $h, \gamma, \beta | y, \mu^*, \phi^*$. Además, la f.d.p. condicional completa de (σ^2, Ω) no se puede promediar directamente. En Chib (1995) se propone una solución para lidiar con esta complicación, la cual es continuar el muestreo de Gibbs para G iteraciones adicionales con las densidades condicionales completas:

$$\begin{aligned} p(\sigma^2, \Omega | y, h, \gamma, \mu^*, \phi^*, \beta) &= p(\Omega|\gamma) p(\sigma^2|h, \mu^*, \phi^*), \\ p(\beta | y, h, \gamma, \mu^*, \phi^*, \sigma^2, \Omega) &= p(\beta|y, h, \mu^*, \phi^*, \sigma^2), \\ p(\gamma | y, h, \mu^*, \phi^*, \beta, \sigma^2, \Omega) &= p(\gamma|y, h, \Omega) \text{ y} \\ p(h | y, \gamma, \mu^*, \phi^*, \beta, \sigma^2, \Omega) &= p(h|y, \gamma, \mu^*, \phi^*, \beta, \sigma^2). \end{aligned} \quad (4.7.1.15)$$

En cada una de estas densidades, μ y ϕ se igualan a μ^* y ϕ^* , respectivamente. Se puede verificar de la teoría de MCMC que las simulaciones $\{h^{(j)}, \gamma^{(j)}, \beta^{(j)}\}$ obtenidas del muestreo de Gibbs utilizando (4.7.1.15) siguen una distribución $h, \gamma, \beta|y, \mu^*, \phi^*$, como se requería (enfoque tomado de Chib (1995), página 1315), consecuentemente:

$$\hat{p}(\sigma^{2*}, \Omega^*) = \frac{1}{G} \sum_{g=1}^G p(\sigma^{2*}, \Omega^*|y, h^{(g)}, \gamma^{(g)}, \mu^*, \phi^*, \beta^{(g)}) \quad (4.7.1.16)$$

es una estimación consistente de (4.7.1.11) donde las simulaciones $\{h^{(g)}, \gamma^{(g)}, \beta^{(g)}\}$ se generan por $p(h, \gamma, \beta|y, \mu^*, \phi^*)$. Este procedimiento conlleva a un incremento del número de iteraciones; aunque no requiere de nueva programación, por lo tanto es fácil de implementarlo.

Ahora, solo falta por estimar $p(\beta^*|y, \mu^*, \phi^*, \sigma^{2*}, \Omega^*)$ en (4.7.1.12); pero nuevamente se tiene que las simulaciones de (h, γ) obtenidas por medio del muestreo de Gibbs utilizando (4.7.1.1), (4.7.1.2), (4.7.1.3), (4.7.1.4) y (4.7.1.5) siguen una distribución $h, \gamma|y$, y no de $h, \gamma|y, \mu^*, \phi^*, \sigma^{2*}, \Omega^*$, como se requiere en (4.7.1.12). Para superar este problema, se continúa con el muestreo de Gibbs con G iteraciones adicionales con las densidades condicionales completas:

$$\begin{aligned} p(\beta|y, h, \gamma, \mu^*, \phi^*, \sigma^{2*}, \Omega^*) &= p(\beta|y, h, \mu^*, \phi^*, \sigma^{2*}), \\ p(\gamma|y, h, \mu^*, \phi^*, \beta, \sigma^{2*}, \Omega^*) &= p(\gamma|y, h, \Omega^*) \text{ y} \\ p(h|y, \gamma, \mu^*, \phi^*, \beta, \sigma^{2*}, \Omega^*) &= p(h|y, \gamma, \mu^*, \phi^*, \beta^*, \sigma^{2*}). \end{aligned} \quad (4.7.1.17)$$

donde en cada una de las densidades, se evalúan μ, ϕ, σ^2 y Ω en $\mu^*, \phi^*, \sigma^{2*}$ y Ω^* , respectivamente. Ahora, se tiene que las simulaciones $\{h^{(g)}, \gamma^{(g)}\}$ del último muestreo siguen una distribución $h, \gamma|y, \mu^*, \phi^*, \sigma^{2*}, \Omega^*$. Consecuentemente, se tiene que:

$$\hat{p}(\beta^*|y, \mu^*, \phi^*, \sigma^{2*}, \Omega^*) = \frac{1}{G} \sum_{g=1}^G p(\beta^*|y, h^{(g)}, \gamma^{(g)}, \mu^*, \phi^*, \sigma^{2*}, \Omega^*) \quad (4.7.1.18)$$

es una estimación consistente de (4.7.1.12).

Finalmente, al substituir (4.7.1.9) en (4.7.1.7) con (4.7.1.13), (4.7.1.16) y (4.7.1.18) en (4.7.1.9), se obtiene la siguiente estimación de la verosimilitud marginal:

$$\begin{aligned} \log \hat{m}(y) &= \log \hat{p}(y|\mu^*, \phi^*, \beta^*, \sigma^{2*}, \Omega^*) + \log p(\mu, \phi, \beta, \sigma^2, \Omega) \\ &\quad - \log \frac{1}{G} \sum_{g=1}^G p(\mu^*, \phi^*|y, h^{(g)}, \gamma^{(g)}, \beta^{(g)}, \sigma^{2(g)}, \Omega^{(g)}) \\ &\quad - \log \frac{1}{G} \sum_{g=1}^G p(\sigma^{2*}, \Omega^*|y, h^{(g)}, \gamma^{(g)}, \mu^*, \phi^*, \beta^{(g)}) \\ &\quad - \log \frac{1}{G} \sum_{g=1}^G p(\beta^*|y, h^{(g)}, \gamma^{(g)}, \mu^*, \phi^*, \sigma^{2*}, \Omega^*), \end{aligned} \quad (4.7.1.19)$$

donde: $\hat{p}(y|\mu^*, \phi^*, \beta^*, \sigma^{2*}, \Omega^*)$ está dado por (4.7.1.8).

4.8 Probabilidades que $\alpha_t \neq 0$

Se asume que su f.d.p. a priori es igualmente probable que la restricción de $\alpha_t = 0$ se mantenga o no, es decir, $P(\alpha_t = 0) = P(\alpha_t \neq 0) = 0,5$. Entonces, la probabilidad a favor de que $\alpha_t = 0$ se puede calcular por medio de la razón de Savage-Dickey:

$$PO_t = \frac{p(\alpha_t = 0|y)}{p(\alpha_t = 0)},$$

donde el numerador de la f.d.p. a posteriori marginal de α_t se evalúa en cero. Primero, la f.d.p. a posteriori marginal $p(\alpha_t|y)$ es desconocida; pero se conoce que $p(\alpha_t|y, h, \Omega)$ es gaussiana; para ver esto, se conoce que la f.d.p. conjunta $p(\gamma|y, h, \Omega)$ es gaussiana (ver ecuación (4.2.0.16)), puesto que todas las f.d.p. marginales de una gaussiana conjunta son gaussianas, así que $p(\gamma|y, h, \Omega)$ lo es. Por consiguiente, se puede utilizar el muestreo de Gibbs para obtener $p(\alpha_t = 0|y)$ al utilizar la media de las simulaciones como estimador de mínimo error cuadrático medio de las simulaciones de $p(\alpha_t = 0|y, h^{(i)}, \Omega^{(i)})$, donde $\{h^{(i)}, \Omega^{(i)}\}$ se simulan con sus f.d.p. a posteriori marginales completas. Finalmente, se puede calcular las probabilidades a posteriori:

$$P(\alpha \neq 0|y) = \frac{1}{1 + PO_t}.$$

4.9 Criterios de comparación entre modelos

En esta sección se brinda un vistazo a varios criterios para comparar el poder predictivo del modelo SVM-TVP con los modelos GARCH, GARCH-2, GARCH-J, GARCH-M, SV, SV-2, SV-J y SV-M. Todos ellos basados en la verosimilitud marginal o integral.

4.9.1 Factor bayesiano, criterio de la información de Schwarz y Akaike

Se supone que se desea comparar el conjunto de modelos $\{M_1, \dots, M_i\}$. Cada modelo M_i es formalmente definido por dos componentes separados: una función de verosimilitud $p(y|\psi_i, M_i)$ que depende del vector de parámetros ψ_i de dimensión p_i que especifica el modelo y la densidad a priori $p(\psi|M_i)$. Un criterio popular de comparación es el factor bayesiano (FB) en favor de M_i y en contra de M_j , se define como:

$$BF_{ij} = \frac{p(y|M_i)}{p(y|M_j)},$$

donde $p(y|M_i)$ es la verosimilitud marginal bajo el modelo M_i , la cual se calculó al final de la anterior sección. Esta verosimilitud marginal se puede interpretar como una densidad de predicción de los datos bajo el modelo M_i evaluado en los datos actuales y . Por lo tanto, si los datos observados son verosímiles (probables) bajo el modelo, la verosimilitud marginal sería más grande. Además, se sigue que $BF_{ij} > 1$ indica que los datos observados son mucho más verosímiles bajo el modelo M_i , comparados al modelo M_j ; y de esta manera, son vistos como una evidencia en favor del modelo M_i .

En segundo lugar, se tiene al criterio de la Información de Schwarz (BIC) para el modelo M_i , el cual se define como:

$$BIC_i = \frac{1}{T} \left[-2 \log p(y|\hat{\psi}_i, M_i) + w \log T \right].$$

El tercer y último criterio es el criterio de Información de Akaike (AIC):

$$AIC = \frac{1}{T} \left[-2T \log p(y|\hat{\psi}_i) + 2w \right],$$

donde: $\hat{\psi}_i$ es la estimación de los parámetros del modelo, T es el tamaño de la muestra y w es la dimensión del vector de parámetros estimados en el modelo. Una propiedad muy útil de este criterio es que compensa el número de parámetros estimados en un modelo, permitiendo así una comparación justa entre modelos que implican diferentes números de parámetros.

Cabe recalcar que el logaritmo del factor bayesiano es asintóticamente equivalente al BIC. Además, el factor bayesiano y el BIC son criterios de selección de modelos consistentes, es decir, ellos seleccionarán asintóticamente el modelo candidato con la correcta estructura con una probabilidad de uno. Por otra parte, el criterio de Akaike no es consistente.

4.10 Pruebas de diagnóstico

El modelo SVM-TVP se basa en un supuesto relativo a los residuos del análisis. Estos residuos deben satisfacer la propiedad de independencia. En esta sección se analiza dos pruebas que se pueden utilizar para establecer si los residuos satisfacen esta suposición. Estas pruebas se aplican a lo que se conoce como errores de predicción estandarizados, los cuales se definen por:

$$e_t = \frac{v_t}{\sqrt{\exp h_t}},$$

donde: v_t son errores de las predicciones a un paso y $\exp h_t$ es la volatilidad del modelo.

La suposición de la independencia de los residuos se puede verificar con el estadístico de Box-Ljung. Al establecer:

$$r_k = \frac{\sum_{t=1}^{n-k} (e_t - \bar{e})(e_{t+k} - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2},$$

denota la autocorrelación residual para k retardos, donde \bar{e} es la media de los n residuos. El estadístico de Box-Ljung, se define por:

$$Q(k) = n(n+2) \sum_{l=1}^k \frac{r_l^2}{n-l},$$

en cambio, el estadístico de Li-McLeod, se define por:

$$Q2(k) = \frac{k(k+1)}{2T} + T \sum_{l=1}^k \frac{r_l^2}{n-l}.$$

En ambos, se rechaza la hipótesis nula (los datos se distribuyen de forma independiente, al nivel del 5 %) cuando:

$$Q(k) > \chi_{(k-w;0,05)}.$$

Existen varios criterios para elegir el número de retardos (k) que se utilizan para las pruebas, por ejemplo: el más común es $\min(20, T-1)$, donde T es la longitud de la serie. Este criterio es propuesto en la página 314 de Box et al. (1994). El segundo es tomar $k = \log T$ propuesto en la página 33 de Tsay (2005).

Parte II

Marco Práctico

CAPÍTULO 5

Resultados computacionales

En el siguiente capítulo se reportan varios resultados de la estimación de los parámetros del modelo SVM-TVP, utilizando las técnicas anteriormente expuestas en el Capítulo 4. Además, se realizan algunas pruebas de diagnóstico del modelo y se revisan sus predicciones. Por último, se compara el desempeño del modelo SVM-TVP con otros modelos de volatilidad condicional y estocástica, utilizando: la verosimilitud marginal, el factor bayesiano y los criterios de información de Akaike (AIC) y Schwarz (BIC).

El presente trabajo utiliza los precios del crudo Oriente obtenidos en la página web de Sistema Nacional de Información (SNI). En esta página web se obtuvieron 2.073 registros con los precios del crudo Oriente desde el 2 de enero de 2012 hasta el 3 de septiembre de 2017. En los resultados computacionales expuestos en este capítulo, se utilizan los precios promedios semanales del crudo Oriente, lo que forma una serie de tiempo de 296 precios promedios semanales. En la primera parte de este capítulo, se utilizarán 284 observaciones para estimar los parámetros del modelo y los restantes 12 se guardarán para compararlos con las predicciones del modelo, y en la segunda parte, se utilizarán 272 para estimar los parámetros del modelo y los restantes 24 para compararlos con las predicciones. Se utilizará la siguiente transformación

$$y_t = 100 \log \left(\frac{z_t}{z_{t-1}} \right), \quad (5.0.0.1)$$

donde: z_t son los precios promedios semanales. En la Figura 5.1 se representa la evolución histórica de las series y_t y z_t . Para volver a la serie original, se utiliza la siguiente fórmula

$$z_t = z_{t-1} \exp(y_t/100). \quad (5.0.0.2)$$

Se utilizó el código de Chan (2017) en MATLAB¹ (ver **Anexo A**) para el modelo de volatilidad estocástica, con parámetros que varían en el tiempo (SVM-TVP), y se utilizó el código empleado en Chan y Grant (2016)² para los modelos: GARCH, GARCH-2, GARCH-J, GARCH-M, SV, SV-2, SV-J y SV-M. En el Cuadro 5.1 se exponen los tiempos de

¹Disponible en <http://joshuachan.org/code/SVM.zip>

²Disponible en http://joshuachan.org/code/code_GARCH.SV.zip

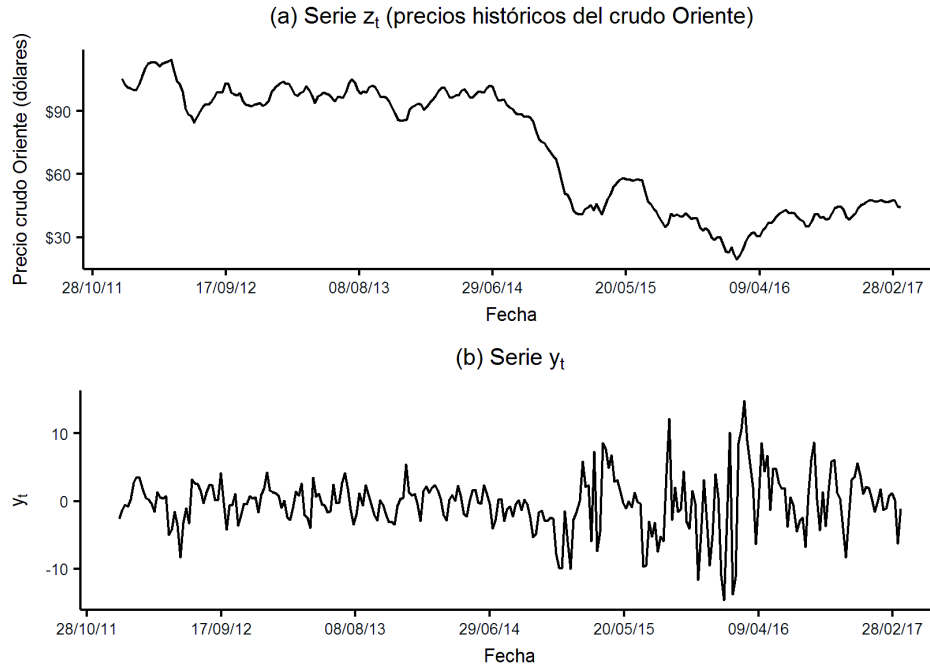


Figura 5.1 Evolución histórica desde el 2012 a 2017 de (a) la serie z_t y (b) la serie y_t .

ejecución de los códigos utilizando el software matemático MATLAB 2013a en una laptop Acer 5552-5898 con procesador AMD Phenom II X3 N830 de tres núcleos con 4 GB de memoria RAM DDR3.

Cuadro 5.1: Tiempos de ejecución de los códigos para los diferentes modelos en Matlab.

Modelo	Tiempo (segundos)
GARCH	380,90
GARCH-2	634,94
GARCH-J	1387,48
GARCH-M	278,76
SV	146,97
SV-2	153,87
SV-J	1.098,75
SV-M	154,31
TVP-SVM	1.255,04
TVP-SVM con $\alpha = 0$	607,47

Los resultados están basados en 10.000 simulaciones de la densidad a posteriori y la verosimilitud marginal. En la Figura 5.2 se pueden observar las primeras 5.000 iteraciones. Al observar (b), (c), (d), (e), (f), (g) y (h) de la Figura 5.2, se puede concluir que después de 1.500 iteraciones se empiezan a estabilizar las estimaciones; por esta razón se descartan las primeras 2.000 (generalmente, se acostumbra a descartar las primeras 2 000 iteraciones (ver Kim et al. (1998))), a esta práctica se llama periodo de quemado, el cual es un periodo en el que las simulaciones obtenidas dependerán del valor tomado como estado inicial de

las cadenas de Márkov, es decir, que después de este periodo de transición, las simulaciones obtenidas no dependerán de los valores iniciales tomados. Se podría evitar la práctica del período de quemado si se eligieran valores iniciales que estén en una zona de alta densidad. Por ejemplo, $\mu = \log(\text{Var}(y))$ (donde $\text{Var}(\cdot)$ denota a la varianza) empieza en una zona de alta densidad, β también podría empezar en una zona de alta probabilidad si se iniciara en $-0,04$ y no en $0,04$.

5.1 Resultados del modelo SVM-TVP con pronósticos a 12 semanas

La serie de tiempo posee 296 precios promedios semanales, de los cuales 284 se utilizarán para estimar los parámetros del modelo y los restantes 12 se guardarán para compararlos con las predicciones del modelo SVM-TVP, y los otros modelos expuestos en la sección 1.3 de este trabajo de titulación.

El muestreo MCMC fue inicializado al establecer: $\beta = 0$, $\phi = 0,98$, $\sigma^2 = 0,024^2$, $\Omega = \text{diagonal}(0,2^2; 0,5^2)$ y $\mu = \log(\text{Var}(y))$. Por último, se inicializa h_t con una normal con media μ y varianza σ^2 . Para la extracción de las simulaciones de las distribuciones a posteriori de los parámetros y estados, se realizan 10.000 iteraciones, de las cuales las 2.000 se desechan (generalmente se acostumbra en la bibliografía a descartar las primeras 2.000 iteraciones (ver Kim et al. (1998))). Al establecer el error cuadrático medio (ECM) como función de riesgo, la estimación de Bayes es la media de la distribución a posteriori de las 8.000 iteraciones restantes. Después, se calcula la probabilidad dinámica que $\alpha_t \neq 0$ y la verosimilitud marginal con el mismo número de iteraciones. Por último, se calculan los puntos de predicción, utilizando una vez más a la media de las simulaciones como estimador de mínimo error cuadrático medio (MECM).

Cabe recalcar que se utilizó el muestreo de Gibbs para extraer las simulaciones de las distribuciones a posteriori de los parámetros (excepto μ y ϕ); pero para extraer las simulaciones de la f.d.p. a posteriori de la volatilidad en escala logarítmica (h_t), se utilizó algoritmo de ARMH con función generadora a una aproximación gaussiana con una tasa de aceptación del 99,97%; en cambio, para extraer simultáneamente las simulaciones de la f.d.p. a posteriori de μ y ϕ , se utilizó el algoritmo de Metropolis-Hasting (M-H) con f.d.p. generadora a una t de Student bivalente con una tasa de aceptación del 65,55%, lo cual no es tan alta como en el muestreo de h_t ; pero aún continua siendo aceptable.

5.1.1 Hiperparámetros de las distribuciones a priori

En esta sección, se proveen los detalles de los hiperparámetros utilizados en el modelo. Las distribuciones a priori están descritas en (1.3.3.4) en la subsección 1.3.3. En particular, μ , ϕ y β siguen distribuciones normales, con medias: $\mu_0 = 1,968$, $\phi_0 = 0,5$ y $\beta_0 = -0,03$, y varianzas: $V_\mu = 0,05$, $V_\phi = 0,1$ y $V_\beta = 0,15^2$, respectivamente.

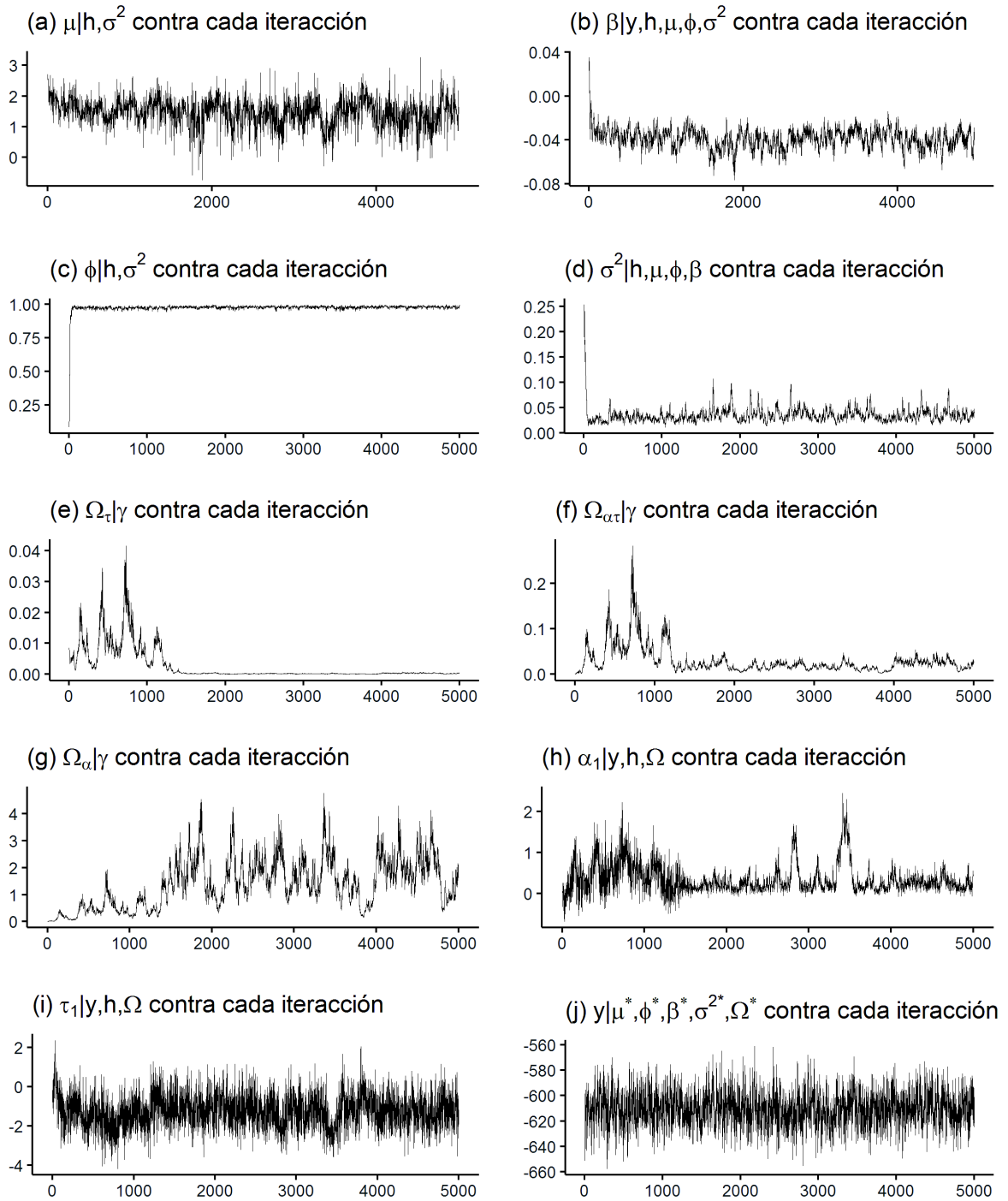


Figura 5.2 Primeras 5.000 simulaciones de (a)-(g) las f.d.p. a posteriori de los parámetros μ , β , ϕ , σ^2 y Ω , (h)-(i) los estados α y τ para el primer intervalo de tiempo después de la inicialización, y (j) la verosimilitud en un punto de alta densidad.

Las varianzas de los estados h_t y γ_t son independientemente distribuidas, y tienen las siguientes distribuciones a priori: $\sigma^2 \sim \mathcal{IG}(v_{\sigma^2}, S_{\sigma^2})$ y $\Omega \sim \mathcal{IW}(v_{\Omega}, S_{\Omega})$, con parámetros $v_{\Omega} = 4$ y $v_{\sigma^2} = 10$; para los parámetros de escala, se establece: $S_{\sigma^2} = 0,02^2(v_{\sigma^2} - 1)$ y $S_{\Omega} = (v_{\Omega} + 3)\text{diag}\{0,1^2; 0,1^2\}$. Estas distribuciones a priori fueron elegidas con estos hiperparámetros, con excepción de Ω y β , adoptando la propuesta utilizada para modelar los precios del crudo W.T.I. en Chan y Grant (2016) en los modelos SV y SV-M.

5.1.2 Estimación del modelo SVM-TVP para el crudo Oriente

En el Cuadro 5.2 se exponen las estimaciones de los parámetros utilizando la media de las simulaciones como estimador de mínimo error cuadrático medio (MECM) y los intervalos de credibilidad al 95 %, que son obtenidos utilizando los cuantiles 0,025 y 0,975 de las simulaciones extraídas.

La estimación asociada con los parámetros del proceso AR(1) de h_t es altamente persistente, con media posterior de ϕ igual a 0,969, esto también implica que resta en 97 % el efecto de μ . También se tiene que β , el coeficiente asociado con el efecto del último precio en la volatilidad estocástica es muy pequeño (-0,034). De hecho, su intervalo de credibilidad al nivel del 95 % está muy cerca de cero, indicando que el último precio del crudo tiene un impacto muy limitado en la volatilidad en escala logarítmica.

Al observar el Cuadro 5.2 y la Figura 5.3 (u), se tiene que la estimación de la correlación entre α_t y τ_t es el único parámetro que tiene dentro de su intervalo de credibilidad, al nivel del 95 %, al cero (-0,086 ; 0,261). En la Figura 5.3 también se observa como las autocorrelaciones empiezan a desaparecer y estabilizarse.

Cuadro 5.2 Estimación a posteriori de los momentos y cuantiles de los parámetros del modelo SVM-TVP guardando 12 observaciones.

Parámetros del modelo SVM-TVP		
Parámetros	Media a posteriori	Intervalo de credibilidad al 95 %
μ	1,325	(0,613 ; 1,910)
β	-0,034	(-0,047 ; -0,022)
ϕ	0,969	(0,947 ; 0,999)
σ^2	0,027	(0,016 ; 0,044)
Ω_{α}	0,07	(0,015 ; 0,190)
$\Omega_{\alpha,\tau}$	0,053	(-0,086 ; 0,261)
Ω_{τ}	0,333	(0,015 ; 1,806)

En la Figura 5.4 se muestra la evolución de h_t y de sus parámetros que varían a través del tiempo asociados α_t , así como sus intervalos de credibilidad asociados al 95 % de credibilidad. Las estimaciones de h_t muestran que la volatilidad de los precios del crudo oriente se disparó desde finales de 2014 y se mantuvieron así, al alza, hasta mediados de 2016 cuando empezó a disminuir. Además, se observa el efecto de la volatilidad a través de los parámetros α_t en los precios del crudo Oriente, que al parecer pierden importancia al incrementar la volatilidad; es decir, mientras mayor sea la volatilidad, menor será su efecto en

la predicción en los precios del crudo. Algo similar sucede con los intervalos de credibilidad de α_t , puesto que mientras mayor sea la volatilidad, menor será el ancho de los intervalos de credibilidad de α_t . Por último, cabe recalcar que el punto más alto de volatilidad se obtuvo en el punto más bajo del precio del crudo.

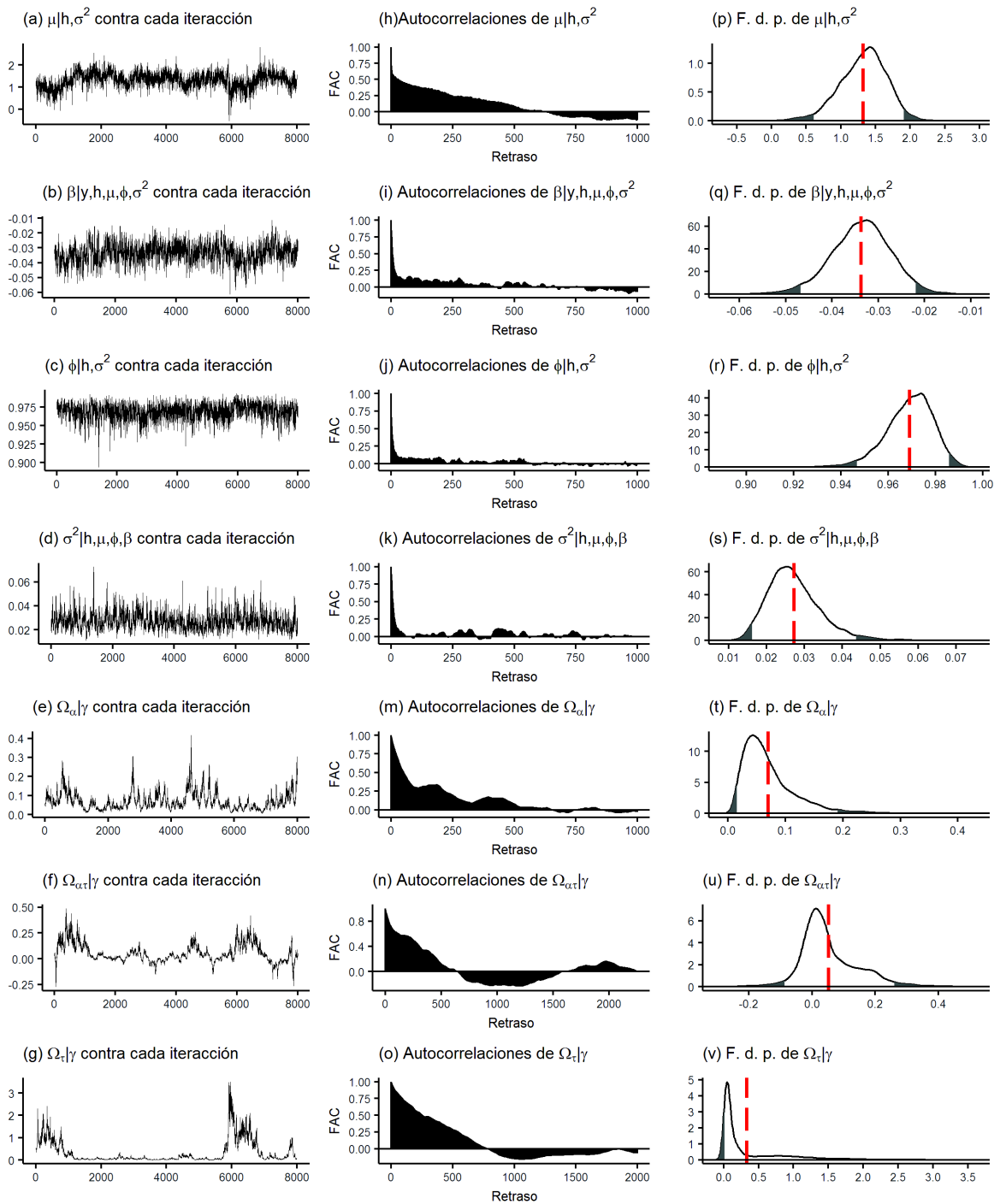


Figura 5.3 (a)-(g) Simulaciones en cada iteración de las f.d.p. a posteriori de los parámetros, (h)-(o) Correlogramas de las simulaciones y (p)-(v) Histogramas de las simulaciones. En total se realizaron 10.000 iteraciones, descartando las primeras 2.000.

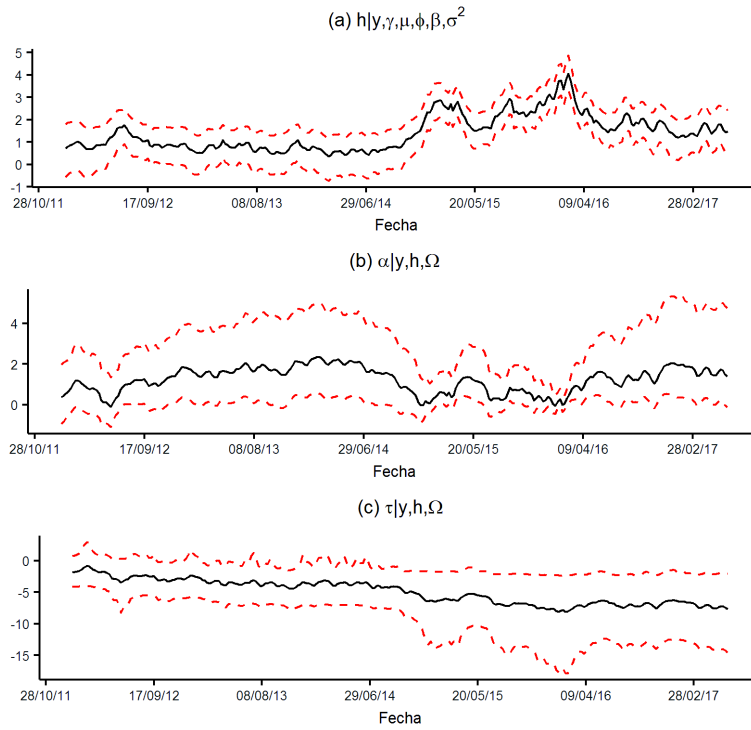


Figura 5.4 (a)-(c) Evolución de las f.d.p. a posteriori de los estados: h_t (a), α_t (b) y τ_t (c). Las medias de las simulaciones se representan con las líneas sólidas negras; y los intervalos de credibilidad, al 95 % de nivel, por líneas punteadas rojas.

5.1.3 Probabilidades que $\alpha_t \neq 0$

Se asume que su f.d.p. a priori es igualmente probable que la restricción de $\alpha_t = 0$ se mantenga o no, es decir, $P(\alpha_t = 0) = P(\alpha_t \neq 0) = 0,5$. La probabilidad a favor de que $\alpha_t = 0$ se calcula por medio de la razón de Savage-Dickey. En la Figura 5.5 se muestran los resultados que indican que la probabilidad de $\alpha_t \neq 0$ es mayor a 0,5 en la mayor parte del tiempo, lo cual justifica su inserción de α_t en el modelo. Además, se observa que existe alta probabilidad que $\alpha_t = 0$, cuando la volatilidad es alta.

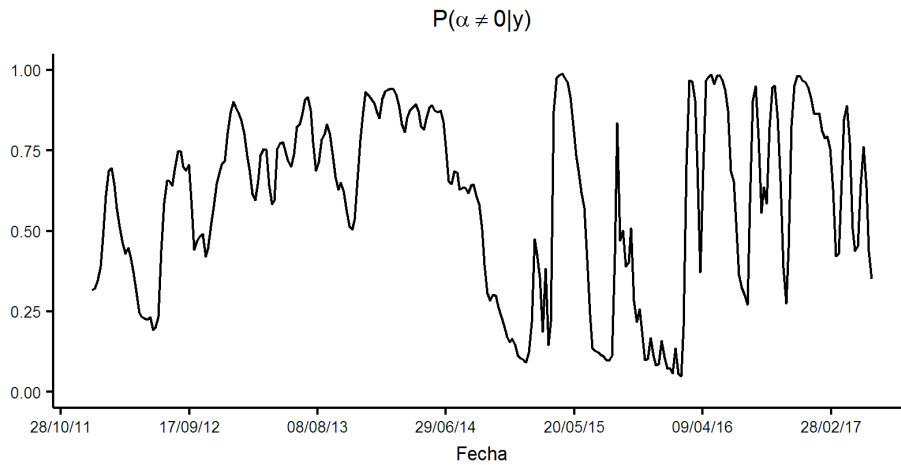


Figura 5.5 Probabilidades a posteriori que $\alpha_t \neq 0$.

5.1.4 Resultados de las pruebas de diagnóstico

Utilizando 283 errores estandarizados, se obtienen los estadísticos de Ljung-Box y Li-McLeod para diferentes retardos con $w = 7$ (ver Figura 5.6). Se observa en (b) de la Figura 5.6 que para todo retardo menor que 30, la hipótesis nula de independencia no es rechazada al nivel del 5 %; pero la prueba de Li-McLeod rechaza la hipótesis nula al nivel de significación del 5 % para todo retardo menor que 16 y al nivel del 10 % para todo retardo menor que 23 (ver (c) de la Figura 5.6). El FAC en (a) de la Figura 5.6 confirma la existencia de una baja correlación.

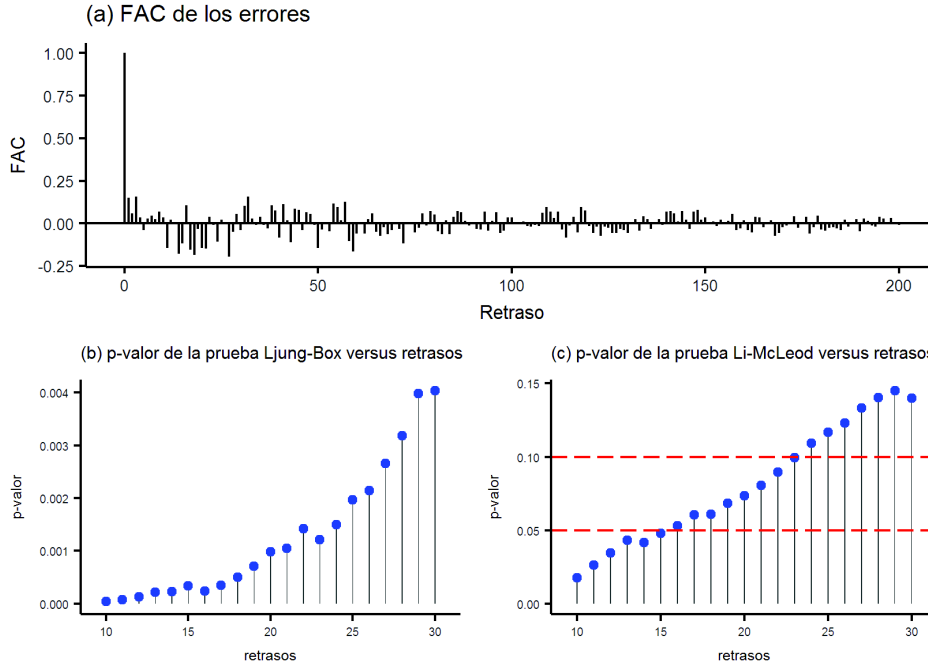


Figura 5.6 (a) Función de auto correlación para los errores, (b) Prueba de Ljung-Box para la independencia de los residuos y (c) Prueba de Li-McLeod para la independencia de los residuos.

5.1.5 Predicciones del modelo SVM-TVP

A continuación, se establecen las predicciones del modelo SVM-TVP para poder evaluar su poder predictivo, al comparar los puntos de predicción con las observaciones guardadas. En la metodología clásica, introducida por Box-Jenkins, se obtiene una estimación de la media de la función de densidad de probabilidad condicional en cada periodo, en el horizonte de predicción; en cambio, en la metodología de MCMC, se generan un par de predicciones: la f.d.p. predictiva y el punto de predicción, que es la media de la f.d.p. predictiva (debido a que se utiliza al error cuadrático como función de riesgo), para cada periodo en el horizonte de predicción.

En la práctica, la media predictiva y la f.d.p. predictiva de y_{t+k} no se pueden calcular analíticamente; en cambio, ellas se obtienen utilizando simulaciones. Más precisamente, en cada iteración de MCMC, dado los parámetros del modelo y estados hasta el tiempo t , se simulan los estados futuros desde el periodo t_0 hasta $t_0 + k - 1$, donde: t_0 es la semana del 09/06/2017 ($t_0 = 284$), utilizando las ecuaciones de estado (ecuaciones (1.3.3.2) y (1.3.3.3), y

simulando los errores ε_t^h y ε_t^γ . También se simulan los errores futuros ε_s^y , para $s = t_0 + 1, \dots, t_0 + k - 1$. Por último, se puede obtener las simulaciones extraídas de la f.d.p. predictiva ($p(y_{t+k}|y_{1:t})$), fácilmente, al reemplazar las simulaciones en (1.3.3.1); y al tomar la media de la f.d.p. predictiva, debido a que se utiliza al error cuadrático como función de riesgo, se estima el punto de predicción ($E\{y_{t+k}|y_{1:t}\}$).

Cabe recalcar que se obtienen las predicciones de z_t (la serie original) de la siguiente manera: primero, después de realizar las 20.000 simulaciones de y_t (al utilizar las ecuaciones (1.3.3.1), (1.3.3.2) y (1.3.3.3)) para cada periodo en el horizonte de predicción, se utiliza la ecuación (5.0.0.2) (inversa de la transformación (5.0.0.1)) para obtener las simulaciones de la serie original z_t ; segundo, se estima la media de las simulaciones de z_t , utilizando su promedio; de esta manera, se obtienen, primeramente, las realizaciones la f.d.p. predictiva ($p(z_{t+k}|z_{1:t})$) y en segundo lugar, su correspondiente estimación de la media ($E\{z_{t+k}|z_{1:t}\}$), para cada periodo en el horizonte de predicción. Por lo tanto, los puntos de predicción del modelo SVM-TVP no están subestimados (menores a las predicciones reales), puesto que la estimación de la media (promedio) se aplica después de la transformación convexa y no antes que la transformación convexa (desigualdad de Jensen en página 37 de Knight (1999)).

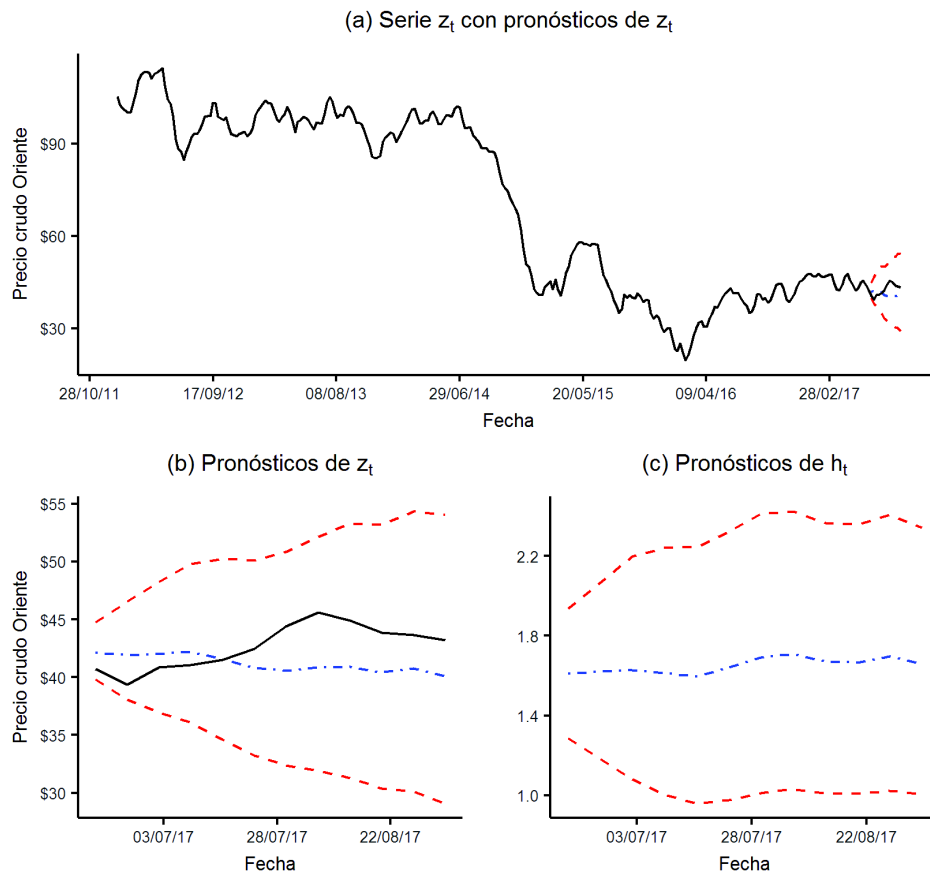


Figura 5.7 (a) Evolución de los precios, 296 registros, en línea negra; evolución de los 12 puntos de predicción (línea azul); y los intervalos de credibilidad, al nivel del 95 % (línea roja). (b) Media de la densidad predictiva como pronósticos para 12 semanas con intervalos de credibilidad, al nivel del 95 %. (c) Evolución de los pronósticos de h_t para 12 semanas con intervalos de credibilidad, al nivel del 95 %.

En la Figura 5.7 y en el Cuadro 5.3, se observa que los puntos de predicción y los

intervalos de credibilidad al 95 %, con $k = 1, \dots, 12$, es decir, los valores esperados y los intervalos de credibilidad, al nivel del 95 %, que se obtienen utilizando los cuantiles 0,025 y 0,975 de las simulaciones extraídas, para los precios del crudo Oriente para 12 semanas. Se puede observar que ninguno de los 12 precios promedios semanales está fuera del intervalo de credibilidad, al nivel del 95 %. También se tiene que los intervalos de credibilidad van creciendo, según aumenta el número de predicciones k .

La métrica para evaluar los puntos de predicción es la raíz cuadrada del error cuadrático de predicción (RECPM), utilizando la media como punto de predicción, se define por:

$$RECPM = \sqrt{\frac{\sum_{t=t_0}^{T-k} (y_{t+k}^o - E\{y_{t+k}|y_{1:t}\})^2}{T - k - t_0 + 1}},$$

donde: y_{t+k}^o denota los valores observados de y_{t+k} .

Cuadro 5.3: Predicciones e intervalos de credibilidad, al nivel del 95 %, para 12 semanas.

k	Fecha	Observado (z_t)	Predicción	Límite inferior	Límite superior
1	18/06/2017	40,73	42,11	39,81	44,76
2	25/06/2017	39,37	41,91	38,06	46,56
3	02/07/2017	40,88	42,02	36,98	48,27
4	09/07/2017	41,07	42,21	36,05	49,77
5	16/07/2017	41,52	41,58	34,62	50,24
6	23/07/2017	42,47	40,79	33,22	50,07
7	30/07/2017	44,42	40,54	32,36	50,82
8	06/08/2017	45,59	40,86	31,92	52,15
9	13/08/2017	44,89	40,92	31,3	53,29
10	20/08/2017	43,87	40,4	30,36	53,19
11	27/08/2017	43,66	40,76	30,12	54,36
12	03/09/2017	43,21	40,1	29,08	54,05

Una ventaja de este enfoque es que se puede medir y predecir la volatilidad. En la Figura 5.4 se observa como desde finales del primer trimestre del año 2016, la volatilidad del crudo Oriente comenzó a bajar y terminó por estabilizarse. Las predicciones dicen que la volatilidad se estabilizaría en el tercer trimestre de 2017, como efectivamente sucedió (ver la Figura 5.7 (c)).

5.1.6 Modelo SVM-TVP con $\alpha_t = 0$

A continuación, se muestran los resultados para un caso particular del modelo SVM-TVP cuando fijamos $\alpha_t = 0$. De esta manera, este modelo sería un SV, con parámetro que cambia con el tiempo (únicamente τ_t), y no dependería de la volatilidad. Para este objetivo, se utilizó el mismo código (ver Apéndice A) con dos pequeños cambios. Primero, después de extraer las simulaciones de la f.d.p. a posteriori de γ_t ; se fijan a $\alpha_t = 0$. En segundo lugar, después de extraer las simulaciones de la distribución a posteriori de Ω ; se fijan $\Omega_{\alpha\tau}$ y Ω_{α} con un valor muy cercano a cero.

La ventaja de este modelo es que no va a depender de los cambios bruscos de la volatilidad. También se observa que en tiempos de alta volatilidad, el efecto de α_t es muy pequeño como se observó en el modelo SVM-TVP.

Se tienen las siguientes observaciones al comparar los Cuadros 5.4 y 5.2: la primera, se prueba que todos los parámetros son diferentes a cero, al nivel del 95 %, puesto que no se incluye al cero dentro de ninguno de los intervalos de credibilidad de los parámetros; la segunda, β y ϕ no cambian, a diferencia de Ω_τ que aumenta de 0,333 a 1,001.

Cuadro 5.4 Estimación a posteriori de los momentos y cuantiles de los parámetros del modelo SVM-TVP con $\alpha_t = 0$ guardando 12 semanas.

Parámetros	Media a posteriori	Intervalo de credibilidad al 95 %
μ	1,705	(0,773; 2,425)
β	-0,039	(-0,058; -0,022)
ϕ	0,978	(0,960; 0,991)
σ^2	0,036	(0,011; 0,062)
Ω_τ	1,001	(0,092; 2,699)

La independencia de los errores mejora con respecto al modelo SVM-TVP, puesto que en las dos pruebas (Box-Ljung y Li-McLeod) rechazan la hipótesis nula al nivel del 5 %; para todo retardo menor que 30 en la prueba de Box-Ljung y para todo retardo menor que 24 en la prueba de Li-McLeod (ver Figura 5.14).

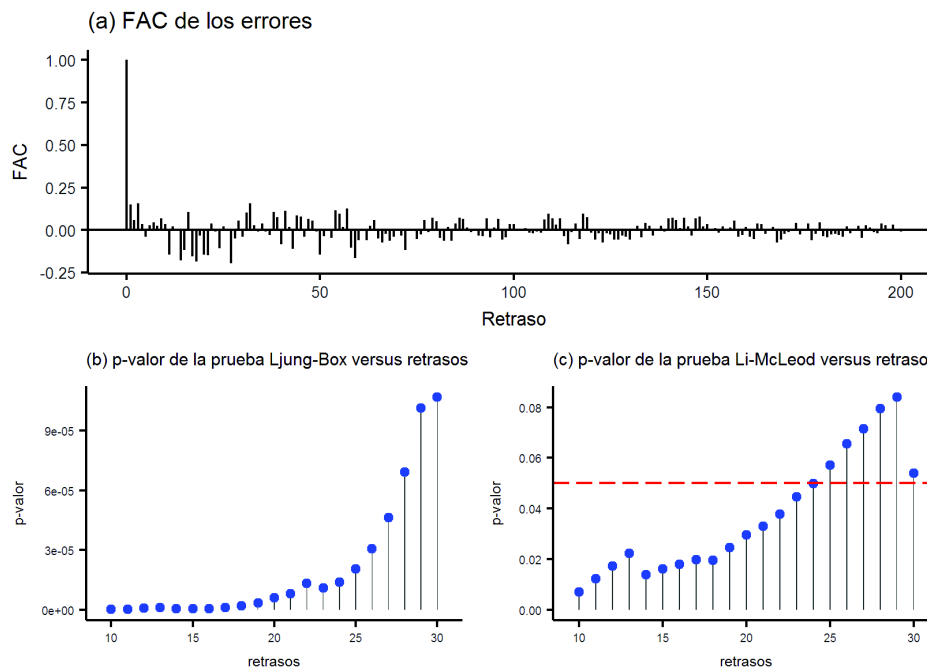


Figura 5.8 (a) Función de auto correlación para los errores del modelo SVM-TVP con $\alpha_t = 0$, (b) Prueba de Ljung-Box para la independencia de los residuos y (c) Prueba de Li-McLeod para la independencia de los residuos.

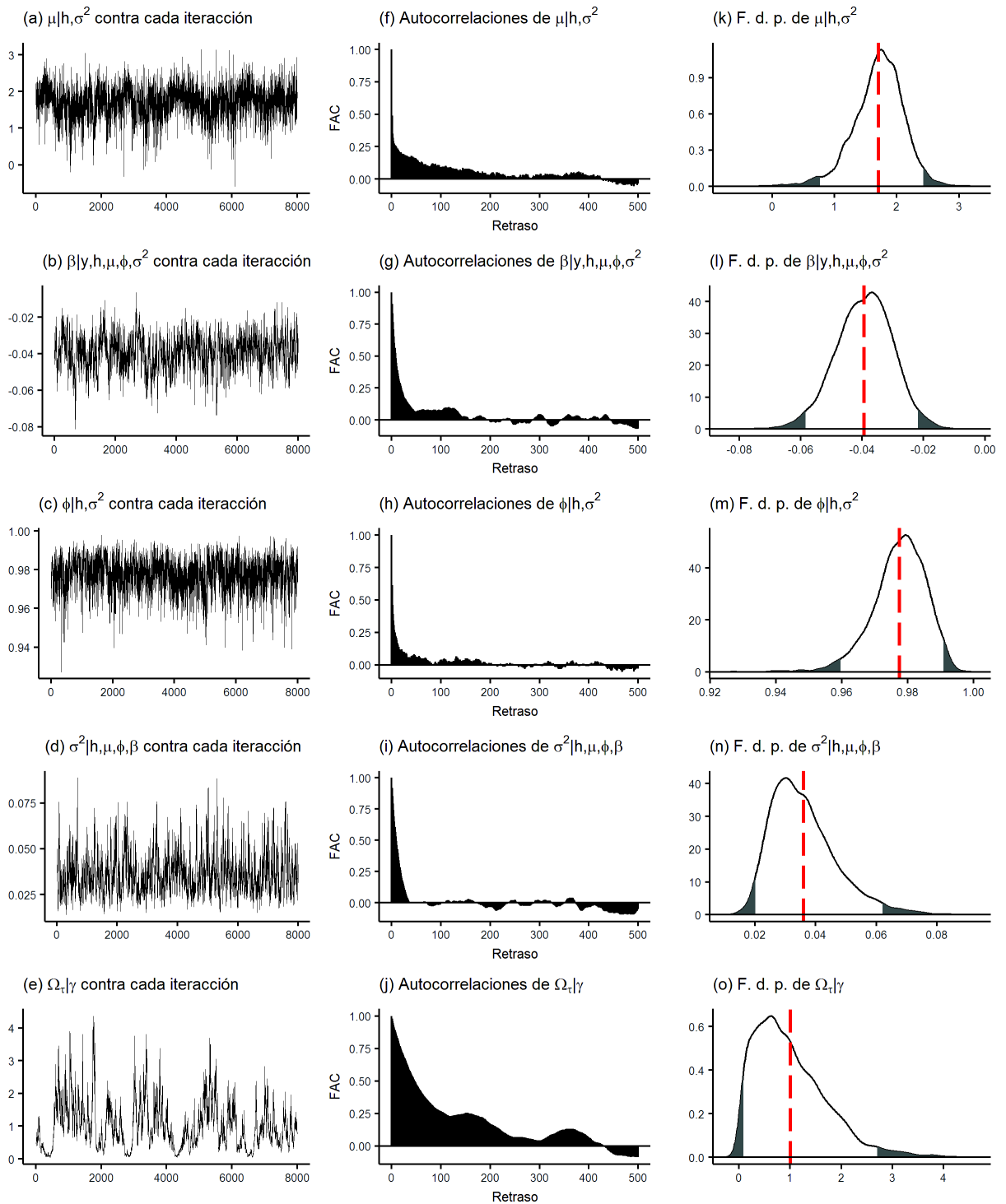


Figura 5.9 (a)-(g) Simulaciones en cada iteración de las f.d.p. a posteriori de los parámetros del modelo SVM-TVP con $\alpha_t = 0$, (h)-(o) Correlogramas de las simulaciones, (p)-(v) Histogramas de las simulaciones. En total se realizaron 10.000 iteraciones, descartando las primeras 2.000.

La gran desventaja del modelo son sus predicciones (ver el Cuadro 5.5 y la Figura 5.10), debido a que su intervalo es más grade que en el modelo original, por ejemplo, para la semana

$k = 12$ es (29, 08; 54, 05), pero para el modelo con $\alpha_t = 0$ es (20, 05; 58, 42). Este último es un problema ya que el precio del crudo nunca ha llegado a estar a 20,05 dólares durante este siglo; de igual manera, sucede para la volatilidad, sus intervalos de credibilidad crecen linealmente a través del tiempo.

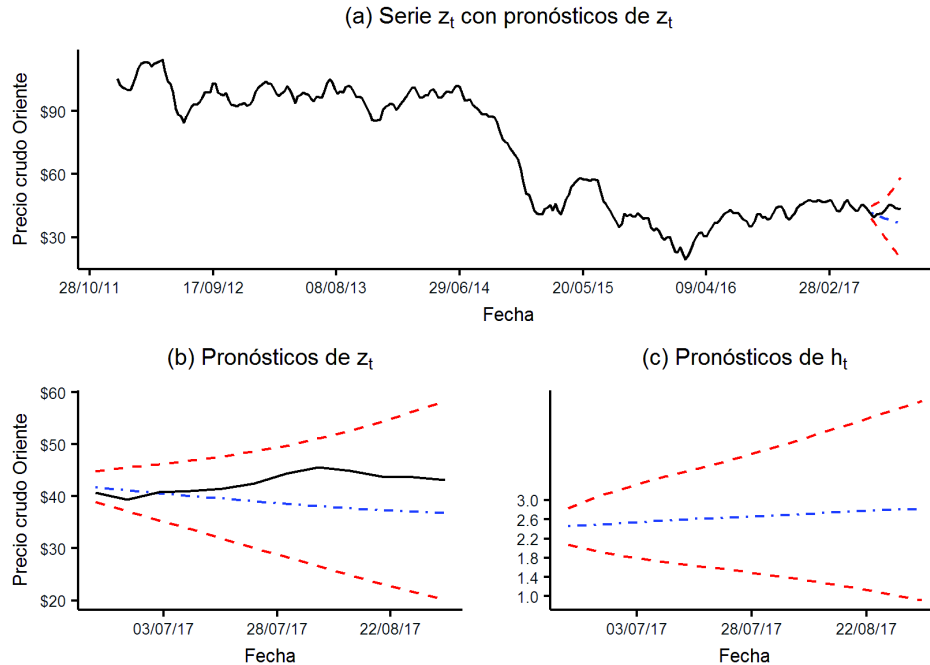


Figura 5.10 (a) Evolución de los precios, 296 registros, en línea negra; evolución de los 12 puntos de predicción (línea azul) del modelo SVM-TVP con $\alpha_t = 0$; y los intervalos de credibilidad, al nivel del 95 % (línea roja). (b) Media de la densidad predictiva como pronósticos para 12 semanas con intervalos de credibilidad, al nivel del 95 %, del modelo SVM-TVP con $\alpha_t = 0$. (c) Evolución de los pronósticos de h_t para 12 semanas con intervalos de credibilidad, al nivel del 95 %, del modelo SVM-TVP con $\alpha_t = 0$.

Cuadro 5.5 Predicciones e intervalos de credibilidad, al nivel del 95 %, del modelo SVM-TVP con $\alpha_t = 0$ para 12 semanas.

k	Fecha	Observado (z_t)	Predicción	Límite inferior	Límite superior
1	18/06/2017	40,73	41,76	38,92	44,82
2	25/06/2017	39,37	41,18	37,14	45,55
3	02/07/2017	40,88	40,61	35,33	46,13
4	09/07/2017	41,07	40,07	33,74	46,94
5	16/07/2017	41,52	39,55	31,86	47,65
6	23/07/2017	42,47	39,06	30,09	48,66
7	30/07/2017	44,42	38,6	28,41	49,66
8	06/08/2017	45,59	38,15	26,59	51,11
9	13/08/2017	44,89	37,75	24,87	52,61
10	20/08/2017	43,87	37,38	23,22	54,42
11	27/08/2017	43,66	37,06	21,6	56,41
12	03/09/2017	43,21	36,78	20,05	58,42

5.1.7 Comparación con otros modelos

Todos modelos: GARCH, GARCH-2, GARCH-J, GARCH-M, SV, SV-2, SV-J y SV-M son estimados utilizando técnicas bayesianas expuestas en Chan y Grant (2016) y se utilizó el código del mismo ³. En el Cuadro 5.6 se muestra como los modelos GARCH y GARCH-M incluyen el cero en al menos en uno de los intervalos de credibilidad de sus parámetros. De igual manera, se puede ver en el Cuadro 5.7, que los modelos de la familia de volatilidad estocástica poseen en algunos parámetros el cero en sus intervalos de credibilidad. Esto indica una clara mala especificación por parte de estos modelos. Por tanto, se llega a la misma conclusión de Chan en Chan y Grant (2016) para los precios del crudo WTI.

Cuadro 5.6 Estimación de los parámetros y sus desviaciones estándar de los modelos GARCH guardando 12 semanas.

Parámetros	GARCH	GARCH-2	GARCH-J	GARCH-M
μ	-0,21 (0,18)	-0,20 (0,18)	-0,05 (0,18)	-0,53 (0,27)
α_0	0,35 (0,17)	0,47 (0,28)	0,67 (0,29)	0,39 (0,19)
α_1	0,13 (0,03)	0,18 (0,04)	0,28 (0,06)	0,13 (0,03)
β_1	0,83 (0,03)	0,43 (0,06)	0,69 (0,09)	0,83 (0,04)
β_2		0,37 (0,07)		
κ			0,08 (0,02)	
μ_κ			-1,04 (0,21)	
σ_κ^2			10,73 (1,54)	
λ				0,03 (0,02)

Cuadro 5.7 Estimación de los parámetros y sus desviaciones estándar de los modelos SV guardando 12 semanas.

Parámetros	SV	SV-2	SV-J	SV-M
μ	-0,16 (0,17)	-0,15 (0,18)	-0,15 (0,19)	0,39 (0,26)
μ_h	2,31 (0,65)	2,23 (0,44)	2,15 (0,80)	2,02 (0,64)
ϕ_h	0,98 (0,01)	1,02 (0,08)	0,98 (0,02)	0,98 (0,01)
ω_h^2	0,05 (0,02)	0,10 (0,05)	0,05 (0,02)	0,04 (0,02)
ρ_h		-0,10 (0,10)		
κ			0,04 (0,03)	
μ_κ			-0,24 (2,31)	
σ_κ^2			10,27 (11,41)	
λ				-0,05 (0,02)

Ahora, se puede comparar los nueve modelos expuestos en la sección 1.3, utilizando: el logaritmo de la función de verosimilitud marginal, el factor bayesiano, AIC, BIC y RECPM ⁴ (ver el Cuadro 5.8). Al utilizar el logaritmo de la función de verosimilitud marginal, se observa que el mejor modelo es SVM-TVP; en segundo y tercer lugar están: SV y SV-2,

³http://joshuachan.org/code/code_GARCH_SV.html

⁴RECPM: Raíz cuadrada del error cuadrático de predicción utilizando la media

respectivamente; en los últimos lugares, se encuentran los modelos SVM-TVP con $\alpha_t = 0$ y SV-M. El logaritmo de la verosimilitud marginal de los modelos SVM-TVP, SV y SV-2 son: $-719,17$, $-752,2$ y $-752,6$, respectivamente; esto implica factores bayesianos de $2,21 \times 10^{14}$ en favor de SVM-TVP en contra de SV y $3,29 \times 10^{14}$ en favor de SVM-TVP en contra de SV-2 (observar el Cuadro 5.8). Una vez más, se muestra que el SVM-TVP es el mejor modelo al utilizar los criterios de información de Akaike y Schwarz; sin embargo, el modelo SVM-TVP con $\alpha_t = 0$ queda en segundo lugar (ver el Cuadro 5.8). Igualmente, al utilizar la RECPM, se tiene que el modelo SVM-TVP tiene mayor poder predictivo que el modelo SVM-TVP con $\alpha_t = 0$.

Cuadro 5.8 Logaritmo de verosimilitud marginal, integral, BIC, AIC, FB, Q, Q2 y RECPM de los modelos: GARCH, SV, GARCH-2, SV-2, GARCH-J, SV-J, GARCH-M, SV-M, SVM-TVP y SVM-TVP con $\alpha_t = 0$.

Modelos	w	$\log p(y \psi^*)$	$\log \hat{m}(y)$	BIC	AIC	FB	Q	Q2	RECPM
GARCH	4	-746,9	-755,8	5,31	5,31	$8,09 \times 10^{15}$	79,73	37,58	
GARCH-2	5	-744,44	-754,4	5,3	5,29	$1,99 \times 10^{15}$	77,99	35,38	
GARCH-J	7	-740,98	-754,1	5,29	5,29	$1,47 \times 10^{15}$	64,32	28,79	
GARCH-M	5	-745,43	-761,5	5,31	5,3	$2,41 \times 10^{18}$	76,81	35,13	
SV	4	-746,36	-752,2	5,31	5,3	$2,21 \times 10^{14}$	86,12	37,87	
SV-2	5	-752,01	-752,6	5,36	5,35	$3,29 \times 10^{14}$	82,65	37,53	
SV-J	7	-746,1	-753,2	5,33	5,32	$6,01 \times 10^{14}$	82,01	34,65	
SV-M	5	-747,4	-762,7	5,33	5,32	$8,03 \times 10^{18}$	84,42	33,01	
SVM-TVP	7	-587,53	-719,17	4,21	4,21	1	34,56	20,97	5,05
SVM-TVP con $\alpha = 0$	5	-711,41	-775,34	5,07	5,06	$2,48 \times 10^{24}$	48,27	24,17	14,62

Como se puede observar en el Cuadro 5.8, los modelos de volatilidad estocástica tienen un mejor rendimiento que los de volatilidad condicional. La excepción a esta regla es GARCH-M que supera a SV-M, puesto que el modelo GARCH-M tiene un mejor rendimiento que el SV-M. Como se mencionó más temprano, bajo los modelos GARCH, la varianza condicional es una función determinista de sus parámetros y datos pasados; en cambio, en los modelos de volatilidad estocástica, el logaritmo de la volatilidad es una variable aleatoria.

Los modelos de volatilidad estocástica son más robustos a la mala especificación y a cambios drásticos de la serie de tiempo. Esto ayuda a explicar ¿por qué estos últimos tienen un mejor rendimiento que sus contra-parte con varianza condicional?. Además, se observa que al insertar una componente de salto al modelo GARCH, esto le da al modelo más flexibilidad en contra de la mala especificación y puntos atípicos, dándole una ventaja frente a los otros modelos GARCH.

5.2 Resultados del modelo SVM-TVP con pronósticos a 24 semanas

La serie de tiempo que posee 296 precios promedios semanales, de los cuales 272 se utilizarán para estimar los parámetros del modelo y los restantes 24 se guardarán para compararlos con las predicciones del modelo SVM-TVP y los otros modelos expuestos en la sección 1.3 de este trabajo de titulación. De igual manera que en la sección anterior (ver

inicio de la sección 5.1), se utilizan los mismos valores para inicializar el muestreo MCMC y se realizan 10.000 iteraciones de las cuales las 2.000 se desechan.

5.2.1 Estimación del modelo SVM-TVP para el crudo Oriente

En el Cuadro 5.9 se exponen las estimaciones de los parámetros, como la media de las simulaciones y los intervalos de credibilidad, al nivel del 95 %, que se obtienen utilizando los cuantiles 0,025 y 0,975 de las simulaciones extraídas.

Cuadro 5.9 Estimación a posteriori de los momentos y cuantiles de los parámetros del modelo SVM-TVP guardando 24 observaciones.

Parámetros	Media a posteriori	Intervalo de credibilidad al 95 %
μ	1,414	(0,765 ; 1,972)
β	-0,032	(-0,045 ; -0,021)
ϕ	0,971	(0,949 ; 0,986)
σ^2	0,027	(0,016 ; 0,043)
Ω_α	0,057	(0,013 ; 0,183)
$\Omega_{\alpha,\tau}$	0,024	(-0,129 ; 0,186)
Ω_τ	0,241	(0,008 ; 1,398)

La estimación asociada con los parámetros del proceso AR(1) de h_t es altamente persistente, con media posterior de ϕ igual a 0,971; esto también implica que resta en 97 % el efecto de μ . También se tiene que β , el coeficiente asociado con el efecto del último precio en la volatilidad estocástica, es muy pequeño (-0,032). De hecho, su intervalo de credibilidad, al nivel del 95 %, está muy cerca de cero, indicando que el último precio del crudo tiene un impacto muy limitado en la volatilidad en escala logarítmica. Al observar el Cuadro 5.9, se tiene que la estimación de la correlación entre α_t y τ_t ($\Omega_{\alpha,\tau}$) es el único parámetro que tiene dentro de su intervalo de credibilidad, al nivel del 95 %, al cero.

Al comparar los Cuadros: 5.2 y 5.9, se observa que los parámetros que más varían son: Ω_τ de 0,333 (guardando 12) a 0,241 (guardando 24) y μ de 1,325 (guardando 12) a 1,414 (guardando 24); en ambos se muestra una diferencia de alrededor de una décima; en cambio, los parámetros β , ϕ , σ^2 , Ω_α y $\Omega_{\alpha,\tau}$ varían en máximo tres centésimas. La evolución de h_t , y de los parámetros, que varían a través del tiempo, asociados: α_t y τ_t , y sus intervalos de credibilidad asociados al 95 % de credibilidad son idénticos, como se observa en la Figura 5.4.

En la Figura 5.11 se muestra la evolución de la probabilidad de que $\alpha_t \neq 0$, a través del tiempo. También se puede observar que la probabilidad de que $\alpha_t \neq 0$ es mayor a 0,5 durante mayor parte del periodo, lo cual justifica la inserción de α_t en el modelo SVM-TVP. Además, se observa que existe alta probabilidad de que $\alpha_t = 0$ cuando la volatilidad es alta.

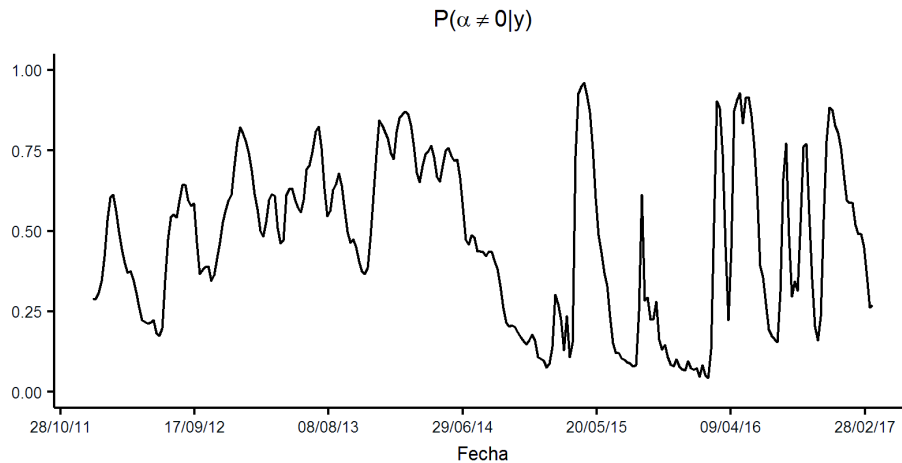


Figura 5.11 Probabilidades a posteriori que $\alpha_t \neq 0$.

5.2.2 Resultados de las pruebas de diagnóstico

Utilizando 283 errores estandarizados se obtienen los estadísticos de Ljung-Box y Li-McLeod para diferentes retardos con $w = 7$ (ver Figura 5.12). Se observa en (b) de la Figura 5.12 que para cualquier retardo menor a 30, la hipótesis nula de independencia no es rechazada al nivel del 5%; pero para la prueba de Li-McLeod se rechaza la hipótesis nula al nivel del 5% para todo retardo menor a 15 y al nivel del 10% para todo retardo menor a 23 (ver (c) de la Figura 5.12). La FAC en (a) de la Figura 5.12 confirma la existencia de una baja correlación.

5.2.3 Predicciones del modelo SVM-TVP

Las predicciones se evalúan para $t = t_0, \dots, t_0 + k - 1$, donde: t_0 es la semana del 19/03/2017 ($t_0 = 272$) y $k = 1, \dots, 24$ son las semanas a predecir. En la Figura 5.13 y en el Cuadro 5.10, se observan que los puntos de predicción y los intervalos de credibilidad, al nivel del 95%, para los precios del crudo Oriente para 24 semanas. También se puede observar que ningún punto observado se encuentra fuera del intervalo de credibilidad al nivel del 95%. Además, se tiene que los intervalos de credibilidad, al nivel del 95%, van creciendo según aumenta el número de predicciones.

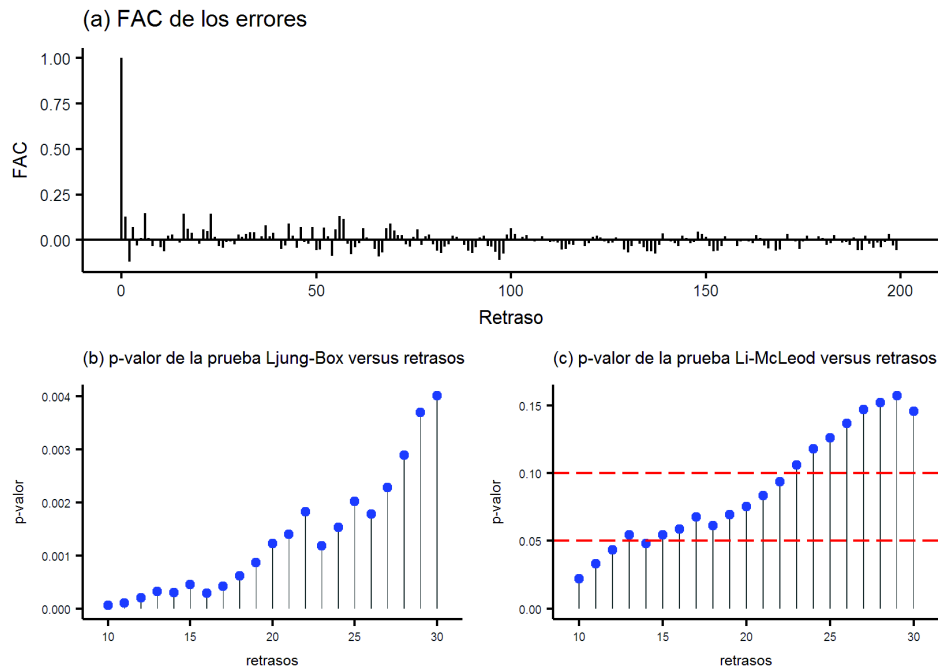


Figura 5.12 (a) Función de auto correlación para los errores, (b) Prueba de Ljung-Box para la independencia de los residuos y (c) Prueba de Li-McLeod para la independencia de los residuos.

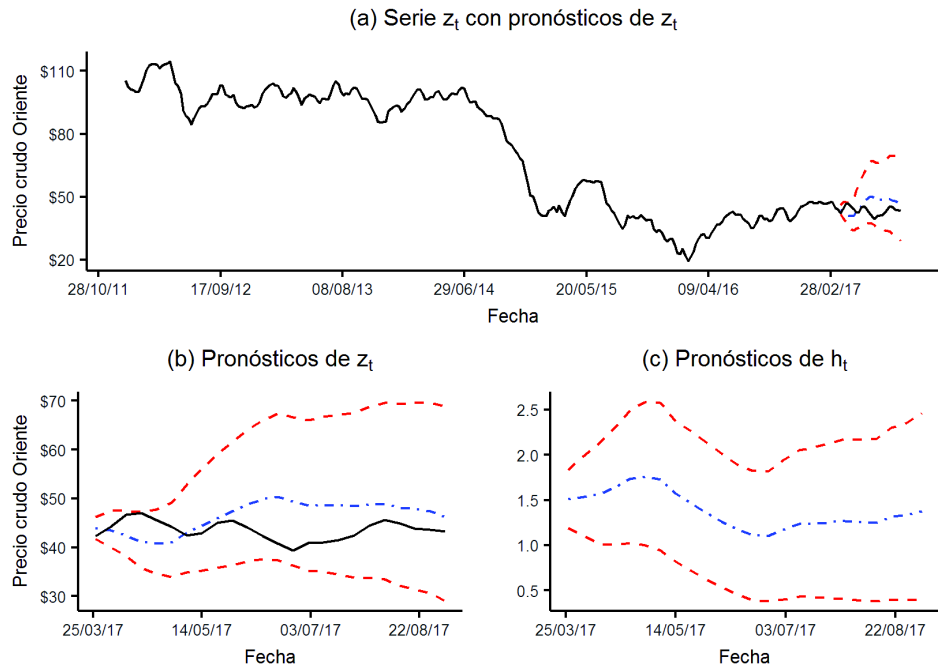


Figura 5.13 (a) Evolución de los precios, en línea negra; evolución de los 24 puntos de predicción, línea azul; y los intervalos de credibilidad, al nivel del 95 %, línea roja. (b) Media de la densidad predictiva como pronósticos para 24 semanas con intervalos de credibilidad, al nivel del 95 %. (c) Evolución de los pronósticos de h_t para 24 semanas con intervalos de credibilidad, al nivel del 95 %.

Cuadro 5.10: Predicciones e intervalos de credibilidad, al nivel del 95 %, para 24 semanas.

k	Fecha	Observado (z_t)	Predicción	Límite inferior	Límite superior
1	26/03/2017	42,26	43,82	41,70	46,19
2	02/04/2017	44,16	43,48	40,11	47,46
3	09/04/2017	46,72	42,46	38,13	47,53
4	16/04/2017	47,01	41,16	35,88	47,23
5	23/04/2017	45,62	40,74	34,61	47,76
6	30/04/2017	44,15	41,03	33,91	49,19
7	07/05/2017	42,41	42,96	34,83	52,85
8	14/05/2017	42,90	44,45	35,27	56,06
9	21/05/2017	45,03	45,94	35,83	59,07
10	28/05/2017	45,48	47,33	36,42	61,55
11	04/06/2017	44,08	48,74	37,16	63,98
12	11/06/2017	42,34	49,86	37,58	66,06
13	18/06/2017	40,73	50,28	37,40	67,27
14	25/06/2017	39,37	49,37	36,34	66,58
15	02/07/2017	40,88	48,53	35,24	66,03
16	09/07/2017	41,07	48,64	35,01	66,76
17	16/07/2017	41,52	48,51	34,45	67,11
18	23/07/2017	42,47	48,41	33,86	67,54
19	30/07/2017	44,42	48,78	33,72	68,71
20	06/08/2017	45,59	48,94	33,47	69,57
21	13/08/2017	44,89	48,14	32,18	69,31
22	20/08/2017	43,87	47,90	31,51	69,56
23	27/08/2017	43,66	47,43	30,61	69,73
24	03/09/2017	43,21	46,18	28,98	68,78

5.2.4 Modelo SVM-TVP con $\alpha_t = 0$

A continuación, se muestran los resultados para un caso particular del modelo SVM-TVP cuando se fija $\alpha_t = 0$. En el Cuadro 5.9 se prueba que todos los parámetros son diferentes a cero, al nivel del 95 %, puesto que no se incluye al cero dentro de ninguno de los intervalos de credibilidad de los parámetros. Al comparar los Cuadros: 5.4 y 5.11, la estimación de los parámetros β y ϕ no cambian; en cambio, μ aumentó de 1,414 a 1,617 y Ω_τ de 0,241 a 1,059.

Cuadro 5.11 Estimación a posteriori de los momentos y cuantiles de los parámetros del modelo SVM-TVP con $\alpha_t = 0$ guardando 24 observaciones.

Parámetros	Media a posteriori	Intervalo de credibilidad al 95 %
μ	1,617	(0,697 ; 2,415)
β	-0,039	(-0,057 ; -0,022)
ϕ	0,978	(0,960 ; 0,991)
σ^2	0,035	(0,020 ; 0,063)
Ω_τ	1,059	(0,014 ; 2,555)

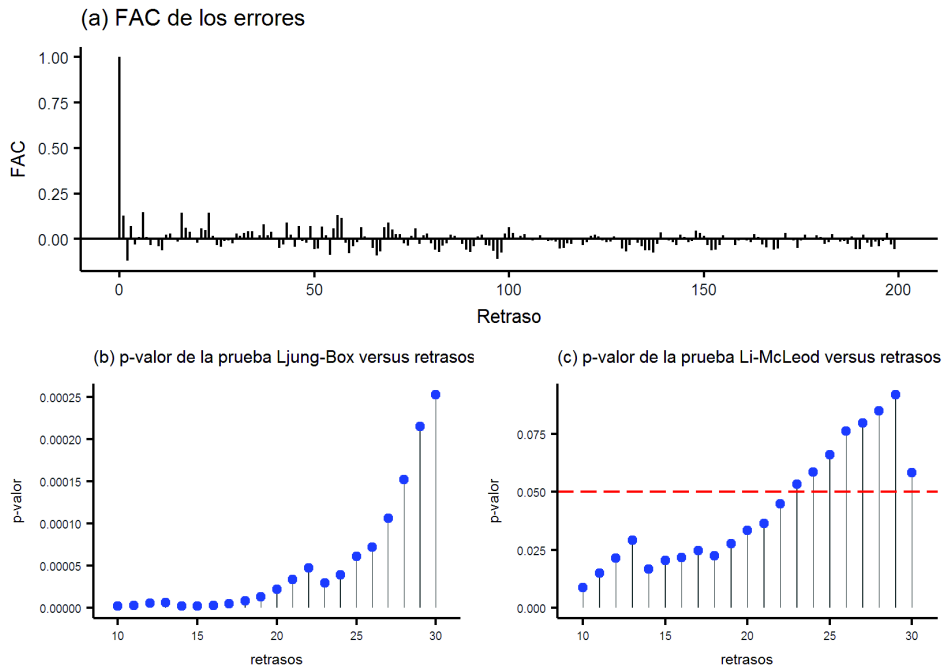


Figura 5.14 (a) Función de auto correlación para los errores del modelo SVM-TVP con $\alpha_t = 0$, (b) Prueba de Ljung-Box para la independencia de los residuos y (c) Prueba de Li-McLeod para la independencia de los residuos.

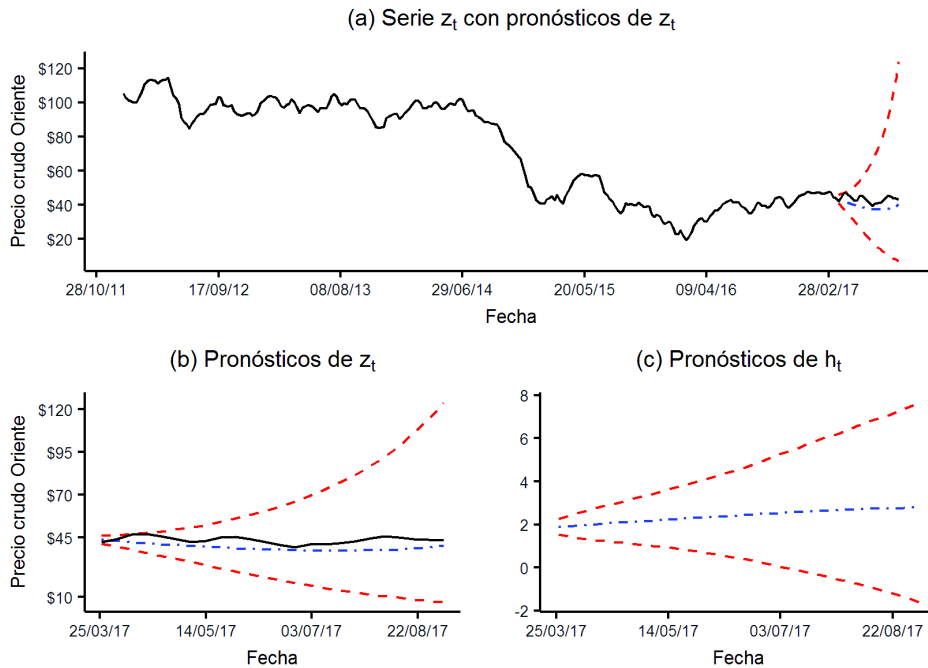


Figura 5.15 (a) Evolución de los precios, 296 registros, en línea negra; evolución de los 24 puntos de predicción (línea azul) del modelo SVM-TVP con $\alpha_t = 0$; y los intervalos de credibilidad, al nivel del 95 % (línea roja). (b) Media de la densidad predictiva como pronósticos para 24 semanas con intervalos de credibilidad, al nivel del 95 %, del modelo SVM-TVP con $\alpha_t = 0$. (c) Evolución de los pronósticos de h_t para 24 semanas con intervalos de credibilidad, al nivel del 95 %, del modelo SVM-TVP con $\alpha_t = 0$.

La independencia de los errores mejora con respecto al modelo SVM-TVP, puesto que

en las dos pruebas (Box-Ljung y Li-McLeod) rechazan la hipótesis nula al nivel del 5 %; para todo retardo menor que 30 en la prueba de Box-Ljung y para todo retardo menor que 23 en la prueba de Li-McLeod (ver Figura 5.14). La gran desventaja del modelo son sus predicciones (ver el Cuadro 5.12 y la Figura 5.15), debido a que su intervalo es más grande que en el modelo original que, por ejemplo, para la semana del 03/09/2017 ($k = 24$) es (28,98 ; 68,78), mientras que para el modelo con $\alpha_t = 0$ es (7,1 ; 124,01).

Cuadro 5.12 Predicciones e intervalos de credibilidad, al nivel del 95 %, del modelo SVM-TVP con $\alpha_t = 0$ para 24 semanas.

k	Fecha	Observado (z_t)	Predicción	Límite inferior	Límite superior
1	26/03/2017	42,26	43,53	41,18	45,9
2	02/04/2017	44,16	42,89	39,45	46,47
3	09/04/2017	46,72	42,27	37,81	46,98
4	16/04/2017	47,01	41,67	36,03	47,68
5	23/04/2017	45,62	41,09	34,23	48,54
6	30/04/2017	44,15	40,55	32,44	49,6
7	07/05/2017	42,41	40,05	30,64	50,91
8	14/05/2017	42,9	39,58	28,87	52,25
9	21/05/2017	45,03	39,14	26,88	54,08
10	28/05/2017	45,48	38,74	25,18	56,13
11	04/06/2017	44,08	38,37	23,29	58,28
12	11/06/2017	42,34	38,04	21,65	60,64
13	18/06/2017	40,73	37,78	19,95	63,02
14	25/06/2017	39,37	37,53	18,33	66,07
15	02/07/2017	40,88	37,36	16,76	69,31
16	09/07/2017	41,07	37,26	15,22	72,98
17	16/07/2017	41,52	37,23	13,96	76,99
18	23/07/2017	42,47	37,27	12,72	81,2
19	30/07/2017	44,42	37,41	11,49	86,61
20	06/08/2017	45,59	37,64	10,53	91,9
21	13/08/2017	44,89	38,01	9,7	98,23
22	20/08/2017	43,87	38,48	8,61	106,28
23	27/08/2017	43,66	39,09	7,9	114,29
24	03/09/2017	43,21	40,29	7,1	124,01

5.2.5 Comparación con otros modelos

Ahora, utilizando las primeras 272 observaciones, se estiman todos modelos: GARCH, GARCH-2, GARCH-J, GARCH-M, SV, SV-2, SV-J y SV-M. Al comparar los Cuadros: 5.6 con 5.13 y 5.7 con 5.14, se muestran pequeños cambios alrededor de una décima en las estimaciones, excepto en μ_κ que decreció de -1,04 a -4,25 en los modelos GARCH (ver Cuadro 5.13) y μ_κ aumentó de 10,73 a 11,07 en los modelos de volatilidad estocástica (ver Cuadro 5.14).

Cuadro 5.13 Estimación de los parámetros y sus desviaciones estándar de los modelos GARCH guardando 24 observaciones.

Parámetros	GARCH	GARCH-2	GARCH-J	GARCH-M
μ	-0,21 (0,18)	-0,19 (0,18)	0,06 (0,19)	-0,54 (0,28)
α_0	0,36 (0,17)	0,38 (0,19)	0,76 (0,27)	0,38 (0,18)
α_1	0,14 (0,03)	0,17 (0,03)	0,26 (0,06)	0,13 (0,03)
β_1	0,84 (0,03)	0,38 (0,09)	0,63 (0,06)	0,84 (0,03)
β_2		0,42 (0,10)		
κ			0,07 (0,02)	
μ_κ			-4,25 (0,99)	
σ_κ^2			11,07 (7,67)	
λ				0,03 (0,02)

Cuadro 5.14 Estimación de los parámetros y sus desviaciones estándar de los modelos SV guardando 24 observaciones.

Parámetros	SV	SV-2	SV-J	SV-M
μ	-0.16 (0.18)	-0.15 (0.19)	-0.16 (0.19)	0.40 (0.26)
μ_h	2.28 (0.74)	2.22 (0.48)	2.10 (0.79)	2.00 (0.64)
ϕ_h	0.98 (0.02)	1.02 (0.08)	0.98 (0.02)	0.97 (0.02)
ω_h^2	0.05 (0.02)	0.11 (0.05)	0.05 (0.02)	0.05 (0.02)
ρ_h		-0.10 (0.10)		
κ			0.04 (0.03)	
μ_κ			0.04 (2.20)	
σ_κ^2			9.38 (9.29)	
λ				-0.05 (0.02)

Ahora, se pueden comparar los nueve modelos expuestos en la sección 1.3, utilizando: el logaritmo de la función de verosimilitud marginal, el factor bayesiano, AIC, BIC y RECPM⁵ (ver el Cuadro 5.15). Al utilizar el logaritmo de la función de verosimilitud marginal, se muestra que el mejor modelo es SVM-TVP; en segundo y tercer lugar están: GARCH-J y SV, respectivamente; en los últimos lugares se encuentran los modelos: SVM-TVP con $\alpha_t = 0$ y SV-M. El logaritmo de la verosimilitud marginal de los modelos: SVM-TVP, GARCH-J y SV son: -710 , $-716,73$ y $-718,92$, respectivamente; esto implica factores bayesianos de $8,37 \times 10^2$ en favor de SVM-TVP en contra de GARCH-J y $7,48 \times 10^3$ en favor de SVM-TVP en contra de SV (observar el Cuadro 5.15). Una vez más, se muestra que el SVM-TVP es el mejor modelo al utilizar los criterios de información de Akaike y Schwarz; pero el modelo SVM-TVP con $\alpha_t = 0$ queda en segundo lugar. Igualmente, al utilizar la RECPM, se concluye que el modelo SVM-TVP tiene mayor poder predictivo que el modelo SVM-TVP con $\alpha_t = 0$.

⁵RECPM: Raíz cuadrada del error cuadrático de predicción utilizando la media

Cuadro 5.15 Logaritmo de verosimilitud marginal, integral, BIC, AIC, FB, Q, Q2 y RECPM de los modelos: GARCH, SV, GARCH-2, SV-2, GARCH-J, SV-J, GARCH-M, SV-M, SVM-TVP y SVM-TVP con $\alpha_t = 0$.

Modelos	w	$\log p(y \psi^*)$	$\log \hat{m}(y)$	BIC	AIC	FB	Q	Q2	RECPM
GARCH	4	-714,8	-723,12	5,09	5,08	$4,99 \times 10^5$	81,99	35,59	
GARCH-2	5	-711,12	-721,63	5,07	5,06	$1,12 \times 10^5$	81,01	34,03	
GARCH-J	7	-705,75	-716,73	5,05	5,04	$8,37 \times 10^2$	64,52	28,4	
GARCH-M	5	-717,16	-728,74	5,11	5,1	$1,38 \times 10^8$	81,73	37,8	
SV	4	-716,64	-718,92	5,1	5,09	$7,48 \times 10^3$	87,88	36,86	
SV-2	5	-714,51	-719,51	5,09	5,08	$1,35 \times 10^4$	84,08	35,24	
SV-J	7	-731,4	-720,02	5,23	5,22	$2,25 \times 10^4$	83,37	33,9	
SV-M	5	-724,6	-729,4	5,16	5,16	$2,66 \times 10^8$	85,36	32,17	
SVM-TVP	7	-578,38	-710	4,15	4,14	1	33,95	20,88	14,13
SVM-TVP con $\alpha = 0$	5	-678,14	-734,71	4,84	4,83	$5,39 \times 10^{10}$	44,86	23,76	20,94

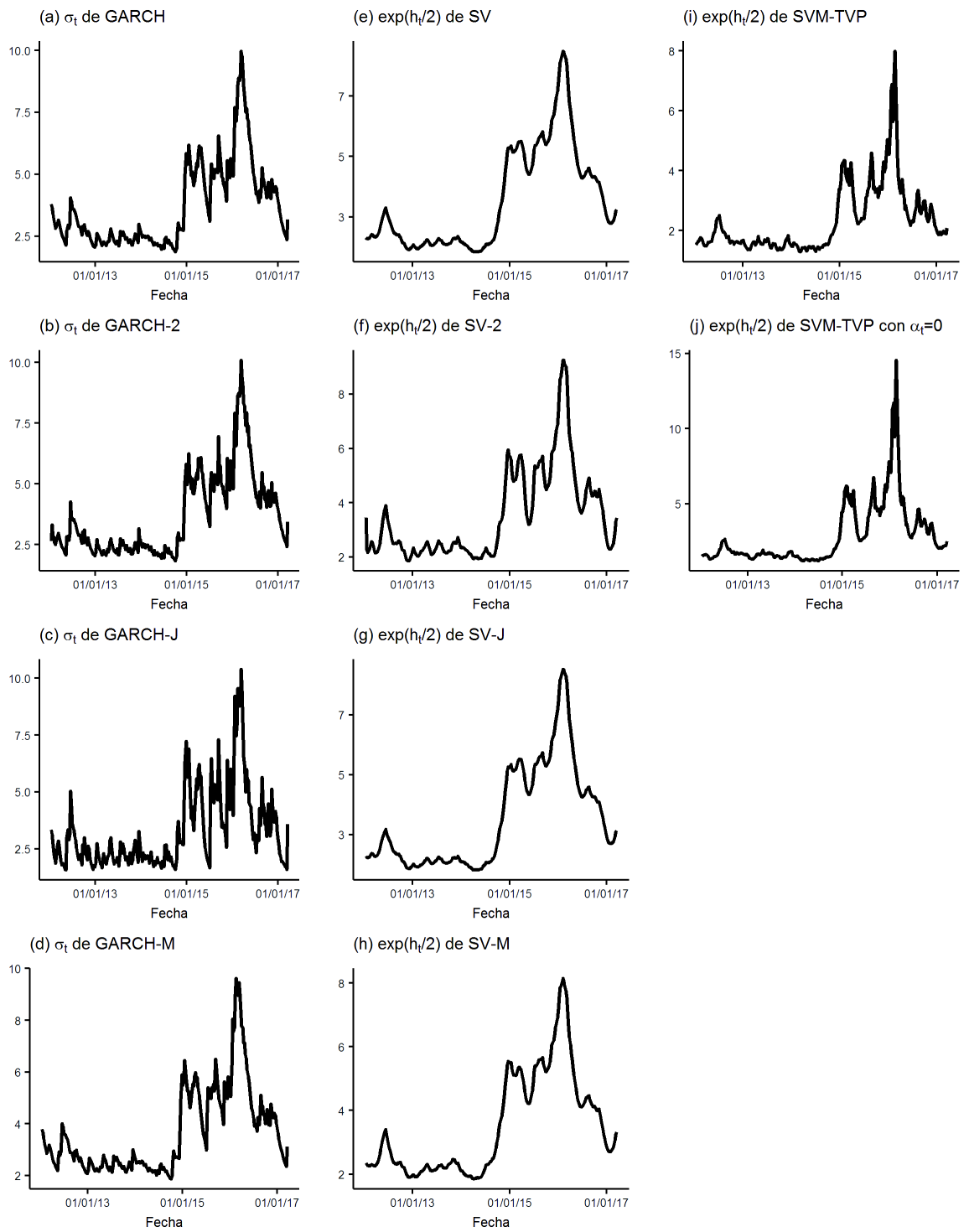


Figura 5.16 Evolución de las desviaciones estándar de los modelos: GARCH, SV, GARCH-2, SV-2, GARCH-J, SV-J, GARCH-M, SV-M, SVM-TVP y SVM-TVP con $\alpha_t = 0$.

CAPÍTULO 6

Conclusiones y recomendaciones

A partir del trabajo realizado, se presentan las siguientes conclusiones y recomendaciones:

1. Como se puede ver en el Cuadro 5.2, el parámetro β del modelo SVM-TVP es muy pequeño y no tiene un impacto significativo en la predicción de los precios del petróleo, es decir, la volatilidad no depende del último precio del crudo.
2. En el modelo SVM-TVP la volatilidad estocástica es inversamente proporcional a los parámetros α_t , es decir, mientras aumenta la volatilidad disminuye el efecto de esta última en la predicción de los precios del petróleo (ver Figura 5.4).
3. En el modelo SVM-TVP se justifica la presencia de $\alpha_t \neq 0$, puesto que sus probabilidades de ser diferentes a cero son en su mayoría mayores a 0,5 (ver Figura 5.5) y porque el poder predictivo de SVM-TVP es mayor al de SVM-TVP con $\alpha_t = 0$ (ver la RECPM en el Cuadro 5.15).
4. El modelo SVM-TVP permite estimar la volatilidad de una manera más precisa, puesto que al ser estocástica utiliza un error adicional, ε^h . Además, es más resistente a la mala especificación, puesto que cuenta con más parámetros (ver Figura 5.16).
5. Como se muestra en la Figura 5.7, el modelo acierta en predecir la estabilidad de la volatilidad del precio del crudo en las primeras semanas de junio a septiembre de 2017; pero la tendencia de los pronósticos indica que los precios se mantendrían estables a finales de 2017, algo que no fue así, puesto que tendieron al alza. Esto muestra lo difícil de predecir los precios de crudo en la actualidad, pues hay un fuerte impacto de orden geopolítico.
6. El Cuadro 5.3 muestra que el modelo SVM-TVP podría ser útil para predecir a un plazo menor a dos meses, puesto que las bandas de los intervalos de credibilidad se hacen cada vez más grandes, mientras mayor sean el número de pronósticos (ver Figura 5.7). Además, existen muchos factores exógenos que influyen en los precios del crudo: la volatilidad macroeconómica y los eventos relacionados con los miembros de la OPEP.

7. El modelo SVM-TVP y el SVM-TVP con $\alpha_t = 0$ fueron los únicos modelos en los cuales sus errores no están correlacionados (ver Figuras 5.6 y 5.8), lo cual muestra que al introducir los parámetros que varían en el tiempo en la ecuación de observación, ayudó a la independencia de los errores.
8. El factor bayesiano (ver Cuadro 5.15) está inflado para este caso, puesto que al estar la verosimilitud marginal muy cerca del cero produce valores muy elevados ($5,39 \times 10^{10}$ para el modelo SVM-TVP en contra del modelo SVM-TVP con $\alpha_t = 0$). Bastaría solo utilizar la verosimilitud marginal para comparar el rendimiento de los modelos, puesto todos los criterios de comparación dependían de este y casi todos muestran las mismas conclusiones (ver Cuadros: 5.15 y 5.8).
9. Utilizando la verosimilitud marginal y los criterios expuestos, se muestra que los modelos estocásticos SV y SV-2 tienen un mejor desempeño que los modelos GARCH y GARCH-2, respectivamente. En cambio, el GARCH-J y GARCH-M se ajustan mejor a los datos observados que los modelos SV-J y SV-M (ver Cuadros 5.8 y 5.15).
10. En futuras investigaciones sería interesante verificar si el modelo SVM-TVP también tiene un buen ajuste para otras series macroeconómicas o financieras. Además, sería interesante pronosticar los precios del crudo Oriente con el modelo GARCH-J con parámetros que varían en el tiempo, debido a que los parámetros que varían en el tiempo podrían ayudar a la independencia de los errores (como sucede con el modelo SV-M y SVM-TVP, ver los Cuadros 5.8 y 5.15) y, de esta manera, el modelo GARCH-J podría ser adecuado.

Referencias bibliográficas

- Anderson, B. y Moore, J. (1979). *Optimal Filtering*. Information and system sciences. Prentice-Hall.
- Berndt, E. R., Hall, B. H., Hall, R. E., y Hausman, J. A. (1974). Estimation and inference in nonlinear structural models. *Annals of Economic and Social Measurement*, 3(4):653–665.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.
- Box, G., Jenkins, G. M., y Reinsel, G. C. (1994). *Time series analysis: forecasting and control*. Prentice Hall, Englewood Cliffs, NJ.
- Castro, A. B. (2008). *Regresión Lineal*. Monografías de Matemáticas y Estadísticas.
- Chan, J. C. (2017). The stochastic volatility in mean model with time-varying parameters: An application to inflation modeling. *Journal of Business & Economic Statistics*, 35.
- Chan, J. C. y Grant, A. L. (2016). Modeling energy price dynamics: Garch versus stochastic volatility. *Energy Economics*, 54:182–189.
- Chan, J. C. y Strachan, R. (2014). The zero lower bound: Implications for modelling the interest rate. (42).
- Chib, S. (1995). Marginal likelihood from the gibbs output. *Journal of the American Statistical Association*, 90(432):1313–1321.
- Chib, S. y Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335.
- Chib, S. y Jeliazkov, I. (2001). Marginal likelihood from the metropolis–hastings output. *Journal of the American Statistical Association*, 96(453):270–281.
- Chib, S. y Jeliazkov, I. (2005). Accept-reject metropolis-hastings sampling and marginal likelihood estimation. *Statistica Neerlandica*, 59(1):30–44.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.
- Engle, R. F., Lilien, D. M., y Robins, R. P. (1987). Estimating time varying risk premia in the term structure: the arch-m model. *Econometrica: Journal of the Econometric Society*, pages 391–407.

- Flores, E. y Merrill, T. (1989). Recent economic performance. *Ecuador: A country study*, pages 107–110.
- Harvey, A., Ruiz, E., y Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2):247–264.
- Jørgensen, B., Lundbye-Christensen, S., Song, P.-K., y Sun, L. (1999). A state space model for multivariate longitudinal count data. *Biometrika*, 86(1):169–181.
- Kim, S., Shephard, N., y Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The Review of Economic Studies*, 65(3):361–393.
- Kitagawa, G. (1987). Non-gaussian state-space modeling of nonstationary time series. *Journal of the American statistical association*, 82(400):1032–1041.
- Knight, K. (1999). *Mathematical Statistics*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis.
- Koopman, S. J. (2002). The stochastic volatility in mean model: empirical evidence from international stock markets. *Journal of applied Econometrics*, 17(6):667–689.
- Kroese, D. P., Chan, J. C., et al. (2014). *Statistical modeling and computation*. Springer.
- Kroese, D. P., Taimre, T., y Botev, Z. I. (2013). *Handbook of monte carlo methods*, volume 706. John Wiley & Sons.
- Lehmann, E. L. y Casella, G. (1998). *Theory of point estimation*. Springer Science & Business Media, 3 edition.
- Nakajima, J. et al. (2011). Time-varying parameter var model with stochastic volatility: An overview of methodology and empirical applications. Technical report, Institute for Monetary and Economic Studies, Bank of Japan.
- Newton, M. A. y Raftery, A. E. (1994). Approximate bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 3–48.
- PetroEcuador (2005). *Hitos de la industria petrolera: 1829-2005*. Relaciones Institucionales de Petroecuador.
- Plante, M. y Traum, N. (2015). Time-varying oil price volatility and macroeconomic aggregates. *Center for Applied Economics and Policy Research*.
- Shumway, R. H. y Stoffer, D. S. (2011). *Time series analysis and its applications: with R examples*. Springer Science & Business Media, tercera edition.
- Tanner, M. A. y Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association*, 82(398):528–540.
- Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical finance*, 4(2):183–204.
- Tierney, L. (1994). Markov chains for exploring posterior distributions. *The Annals of Statistics*, pages 1701–1728.
- Tsay, R. S. (2005). *Analysis of financial time series*. John Wiley & Sons.

APÉNDICE A

Código en Matlab

```
1
2 %=====
3 %% Modelo con parámetros que varían en el tiempo con volatilidad estocástica
4 %%
5 %% y_t = tau_t + alpha_t exp(h_t) + e_t, e_t ~ N(0,exp(h_t)),
6 %% h_t = mu + phi (h_{t-1}-mu) + beta y_{t-1} + v_t, v_t ~ N(0,sig2),
7 %%
8 %% gam_t = (alpha_t, tau_t)'
9 %% gam_t = gam_{t-1} + w_t, w_t ~ N(0,Omega)
10 %%
11 %% Condiciones Iniciales: tau_1 ~ N(0,Vtau), h_1 ~ N(mu,sig2/(1-phi^2)).
12 %%
13 %=====
14
15 clear; clc;
16 nloop = 10000;
17 burnin = 2000;
18 M=200;
19 c=100;
20 is_dym_prob = true; %% true: calcula las probabilidades dinamicas de alpha
21 is_alp_const = false; %% true:
22 load 'PreciosOriente.csv';
23 z=PreciosOriente(1:284,:);
24 y =c*diff(log(PreciosOriente(1:284,:)));
25 zcompleto =PreciosOriente(:,:);
26 T = length(y);
27 tid = linspace(2012,2017.475,T)';
28
29 if is_alp_const && is_alp_const
30     error('Las probabilidades dinámicas TVP-SVM- alp=const no pueden ser calculadas');
31 end
32
33 %% Hiperparámetros a priori
34 phi0 =0; Vphi = 1;
35 mu0 = 2; Vmu = 3;
36 Vgam = 2*eye(2);
37 invVgam = Vgam\speye(2);
38 beta0 = 0; Vbeta = 3^2;
39 nuOmega = 4;
40 nuh =10; Sh = .2^2*(nuh-1);
41 SOmega = (nuOmega+3)*diag([.1^2 .1^2]);
42 model_name = 'UC-SVM';
43
44 %% fdp a priori
45 prior=@(m,l,g,n,k) -0.5*log(2*pi)-0.5*log(Vmu)-0.5*(m-mu0)^2/Vmu ...
46 -0.5*log(2*pi)-0.5*log(Vphi)-0.5*(1-phi0)^2/Vphi ...
47 -0.5*log(2*pi)-0.5*log(Vbeta)-0.5*(g-beta0)^2/Vbeta ...
48 +nuh*log(Sh)-gammaln(nuh)-(nuh+1)*log(n)-Sh/n ...
49 +0.5*nuOmega*log(det(SOmega))-nuOmega*log(2)-gammaln(nuOmega/2) ...
```

```

50     -0.5*(nuOmega+3)*log(det(k))-0.5*log(trace(SOmega*inv(k)));
51
52     %% Inicialización de la cadena de Markov
53     beta = 0.02;
54     mu = log(var(y)); phi = 0; sig2 = .5^2;
55     Omega = diag([.1^2 .1^2]);
56     h = mu + sqrt(sig2)*randn(T,1);
57     exph = exp(h);
58
59     %% Inicialización del almacenaje
60     store_theta = zeros(nloop - burnin,7); %%[mu beta phi sig2 Omega([1 2 4])]
61     store_tau = zeros(nloop - burnin,T);
62     store_alp = zeros(nloop - burnin,T);
63     store_h = zeros(nloop - burnin,T);
64     QQ=zeros(nloop - burnin,30);
65     QQ2=zeros(nloop - burnin,30);
66
67     %% Calcular una pocas cosas a fuera del lazo
68     H = speye(2*T) - sparse(3:2*T,1:2*(T-1),ones(1,2*(T-1)),2*T,2*T);
69     Hphi = speye(T) - sparse(2:T,1:(T-1),phi*ones(1,T-1),T,T);
70     newnuh = T/2 + nuh;
71     counth = 0; countlam = 0;
72     disp(['Comenzando MCMC para ' model_name '.... ']);
73     start_time = clock;
74     for loop = 1:nloop
75
76         %% Muestreo de gamma
77         invOmegagam = [invVgam sparse(2,2*(T-1)); ...
78             sparse(2*(T-1),2) kron(speye(T-1),Omega\speye(2))];
79         Xgam = SURform([exph ones(T,1)]);
80         temp1 = Xgam' * sparse(1:T,1:T,1./exph);
81         Kgam = H' * invOmegagam * H + temp1 * Xgam;
82         Cgam = chol(Kgam, 'lower');
83         gamhat = Kgam \ (temp1 * y);
84         gam = gamhat + Cgam' \ randn(2*T,1);
85         alp = gam(1:2:end);
86         tau = gam(2:2:end);
87
88         %% Muestreo de h
89         HinvsH = Hphi' * sparse(1:T,1:T,[(1-phi^2)/sig2; 1/sig2*ones(T-1,1)]) * Hphi;
90         deltah = Hphi \ ([mu; mu*(1-phi)*ones(T-1,1)] + [0;y(1:end-1)] * beta);
91         HinvsHdeltah = HinvsH * deltah;
92         s2 = (y-tau).^2;
93         errh = 1; ht = h;
94         while errh > 10^(-3)
95             expht = exp(ht);
96             sinvexpht = s2 ./ expht;
97             alp2expht = alp.^2 .* expht;
98             fh = -.5 + .5 * sinvexpht - .5 * alp2expht;
99             Gh = .5 * sinvexpht + .5 * alp2expht;
100            Kh = HinvsH + spdiags(Gh,0,T,T);
101            newht = Kh \ (fh + Gh .* ht + HinvsHdeltah);
102            errh = max(abs(newht - ht));
103            ht = newht;
104        end
105        cholHh = chol(Kh, 'lower');
106        %% Paso - AR:
107        hstar = ht;
108        uh = hstar - deltah;
109        logc = -.5 * uh' * HinvsH * uh - .5 * sum(hstar) + ...
110            - .5 * exp(-hstar) * (y-tau-alp .* exp(hstar)).^2 + log(3);
111        flag = 0;
112        ich=0;
113        while flag == 0
114            ich=ich+1;
115            hc = ht + cholHh \ randn(T,1);
116            vhc = hc - ht;
117            uhc = hc - deltah;
118            alpARc = -.5 * uhc' * HinvsH * uhc - .5 * sum(hc) + ...
119                - .5 * exp(-hc) * (y-tau-alp .* exp(hc)).^2 + .5 * vhc' * Kh * vhc - logc;
120            if alpARc > log(rand)
121                flag = 1;

```

```

122     end
123 end
124 %% Paso - MH
125 vh = h-ht;
126 uh = h-delta;
127 alpAR = -.5*uh'*HinvSH*uh -.5*sum(h) + ...
128     -.5*exp(-h)'*(y-tau-alp.*exp(h)).^2 + .5*vh'*Kh*vh - logc;
129 if alpAR < 0
130     alpMH = 1;
131 elseif alpARc < 0
132     alpMH = - alpAR;
133 else
134     alpMH = alpARc - alpAR;
135 end
136 if alpMH > log(rand) || loop == 1
137     h = hc;
138     expH = exp(h);
139     counth = counth + 1;
140 end
141
142 %% Muestreo de beta
143 ybeta = h(2:end) - mu - phi*(h(1:end)-mu);
144 Dbeta = 1/(1/Vbeta + y(1:end-1)*y(1:end-1)/sig2);
145 betahat = Dbeta*(beta0/Vbeta + y(1:end-1)*ybeta/sig2);
146 beta = betahat + sqrt(Dbeta)*randn;
147
148 %% Muestreo de Omega
149 err = reshape(gam(3:end,:) - gam(1:end-2,:),2,T-1)';
150 Omega = iwishrnd(SOmega + err'*err,nuOmega + T-1);
151
152 %% Muestreo sig2
153 errh = [(h(1)-mu)*sqrt(1-phi^2); h(2:end)-phi*h(1:end-1)-mu*(1-phi)-y(1:end-1)*beta];
154 newSh = Sh + sum(errh.^2)/2;
155 sig2 = 1/gamrnd(newnuh, 1./newSh);
156
157 %% Muestreo conjunto de mu y phi
158 sum1 = sum(h(2:end));
159 sum2 = sum(h(1:end-1));
160 flam = @(x) -(1-x(2))^2/(2*sig2)*(h(1)-x(1))^2 - 1/(2*sig2) * ...
161     sum((h(2:end) - x(2)*h(1:end-1) - x(1)*(1-x(2))-y(1:end-1)*beta).^2) ...
162     - 1/(2*Vphi)*(x(2)-phi0)^2 - 1/(2*Vmu)*(x(1)-mu0)^2;
163 [lamc,g] = proplam(h,sig2,phi0,Vphi,mu0);
164 MHprob = flam(lamc) - flam([mu; phi]) + g([mu; phi]) - g(lamc);
165 if exp(MHprob) > rand
166     mu = lamc(1);
167     phi = lamc(2);
168     countlam = countlam+1;
169 end
170 Hphi = speye(T) - sparse(2:T,1:(T-1),phi*ones(1,T-1),T,T);
171
172 %% Estadístico Q y Q2
173 if loop>burnin
174     i = loop-burnin;
175     store_tau(i,:) = tau';
176     store_h(i,:) = h';
177     store_alp(i,:) = alp';
178     store_theta(i,:) = [mu beta phi sig2 Omega([1 2 4])];
179     %%Cálculo del estadístico Q
180     u = (y-diag(alp)*exp(h)-tau)./exp(h/2);
181     for j = 1:30
182         rtmp = autocorr(u,j);
183         QQ(i,j) = T*((T+2)./(T-(1:j)))*rtmp(2:end).^2;
184     end
185
186     nlag=20;
187     rtmp = autocorr(u,nlag);
188     Q = T*((T+2)./(T-(1:nlag)))*rtmp(2:end).^2;
189
190     %%Cálculo del estadístico Q2 de Li-McLeod
191     u = (y-diag(alp)*exp(h)-tau)./exp(h/2);
192     for j = 1:30
193         rtmp = autocorr(u.^2,j);

```

```

194         QQ2(i, j) = T*((T+2)./(T-(1:j)))*rtmp(2:end).^2;
195     end
196     rtmp = autocorr(u.^2, nlag);
197     Q2 = T*((T+2)./(T-(1:nlag)))*rtmp(2:end).^2;
198     store_Q(i, :) = [Q Q2];
199     logmuphi(i)=-T*log(2*pi)/2-(1-phi)^2*(h(1)-mu)^2/(2*sig2^2)-sum(h(2:T)...
200         -phi*h(1:T-1)-mu*(1-phi))^2/(2*sig2^2)-(phi-phi0)^2/(2*Vphi)...
201         -(mu-mu0)^2/(2*Vmu);
202     end
203     if (mod(loop,1000) == 0)
204         disp([ num2str(loop) ' loops... '])
205     end
206 end
207
208 %% Estimador bayesiano como acción la esperanza y riesgo el error cuadrático
209 tauhat = mean(store_tau)';
210 alphas = mean(store_alp)';
211 hhat = mean(store_h)';
212 thetahat = mean(store_theta)';
213 hCI = quantile(store_h, [.025 .975])';
214 alpCI = quantile(store_alp, [.025 .975])';
215 tauCI = quantile(store_tau, [.025 .975])';
216 thetaCI = quantile(store_theta, [.025 .975])';
217 Q=mean(store_Q);
218 Q2=Q(2);
219 Q=Q(1);
220 Qhat = mean(store_Q)';
221 Qstd = std(store_Q)';
222 yfitted=diag(alp)*exp(h)+tau;
223 ev=diag(alp)*exp(h);
224 e=y-yfitted;
225
226 %% Cálculo de los Pronosticos
227 m=12; %% Predicciones adelante
228 n=20000; %% Simulaciones para las predicciones
229 Ome=[thetahat(5), thetahat(6); thetahat(6), thetahat(7)];
230 store_y = zeros(n,m);
231 store_z = zeros(n,m);
232 store_hh = zeros(n,m);
233 store_ttau=zeros(n,m);
234 store_aalp=zeros(n,m);
235 egamma=randn(n,2)*chol(Ome);
236 store_aalp(:,1)=mean(store_alp(:,end))+egamma(:,1);
237 store_ttau(:,1)=mean(store_tau(:,end))+egamma(:,2);
238
239 eh = normrnd(0, (thetahat(4))^0.5, n, 1);
240 store_hh(:,1)= thetahat(1)+thetahat(3)*(mean(store_h(:,end))-thetahat(1))...
241     +thetahat(2)*y(end)+eh;
242 ey=zeros(n,1);
243 for j=1:n
244     ey(j)=normrnd(0, (exp(store_hh(j,1)))^0.5, 1, 1);
245 end
246 store_y(:,1)=store_ttau(:,1)+ dot(store_aalp(:,1), exp(store_hh(:,1)), 2)+ey;
247 store_z(:,1)=z(end)*exp(store_y(:,1)/c);
248 for i=1:m-1
249     for j=1:n
250         ey(j)=normrnd(0, (exp(store_hh(j,i)))^0.5, 1, 1);
251     end
252     store_y(:,i+1)=store_ttau(:,i)+ dot(store_aalp(:,i), exp(store_hh(:,i)), 2)+ey;
253     store_z(:,i+1)=diag(store_z(:,i))*exp(store_y(:,i+1)/c);
254     a=mvnrnd([0;0],[thetahat(5), thetahat(6); thetahat(6), thetahat(7)], 1);
255     ealp=a(1);
256     etau=a(2);
257     eh = normrnd(0, (thetahat(4))^0.5, n, 1);
258     store_aalp(:,i+1)=store_aalp(:,i)+ealp;
259     store_ttau(:,i+1)=store_ttau(:,i)+etau;
260     store_hh(:,i+1)= thetahat(1)+thetahat(3)*(store_hh(:,i)-thetahat(1))...
261         +thetahat(2)*store_y(:,i)+eh;
262 end
263 yCI = quantile(store_y, [.025 .975])';
264 yFI = median(store_y);
265 hCI = quantile(store_hh, [.025 .975])';

```

```

266 hFIf= median(store_hh);
267
268 yforest=mean(store_z);
269 yLimmits=quantile(store_z,[.025 .975]);
270
271 %% Poder predictivo
272 RMSFE=sqrt((sum(zcompleto(285:end,1)-yforest'))^2/length(zcompleto(285:end,1)));
273
274 %% Gráficos
275 figure;
276 hold on
277     plot(1:12, yforest, '—r', 'LineWidth',2,'Color','blue');
278     plot(1:12, yLimmits, '—r', 'LineWidth',2);
279     plot(1:12, zcompleto(285:end,:), 'LineWidth',2, 'Color','Black');
280     title(['Pronosticos de la precios del crudo'])
281 hold off
282
283
284 figure;
285 hold on
286     plot(1:12, hFIf, 'LineWidth',2,'Color','black');
287     plot(1:12, hCI, '—r', 'LineWidth',2);
288     title(['Pronosticos de la volatilidad'])
289 hold off
290
291 %% Probabilidad de alpha=0
292
293 if is_dym_prob
294     compute_dyn_prob;
295     figure
296     plot(tid,prob,'LineWidth',2,'Color','black');
297     box off; ylim([0 1.02]);
298     title('Probabilidades dinámicas que \alpha_t \neq 0');
299 end
300
301 %% Gráficos
302 figure
303 plot(tid, y); title('y_t'); box off;
304 set(gcf,'Position',[100 100 500 400]);
305
306 figure;
307 subplot(1,2,1);
308 hold on
309     plot(tid, hhat, 'LineWidth',2,'Color','black');
310     plot(tid, hCI, '—r', 'LineWidth',2);
311 hold off
312 title('h_t');
313 subplot(1,2,2);
314 hold on
315     plot(tid, alphas, 'LineWidth',2,'Color','black');
316     plot(tid, alpCI, '—r', 'LineWidth',2);
317 hold off
318 title('\alpha_t');
319 set(gcf,'Position',[100 100 1000 400]);
320
321 figure;
322 subplot(1,2,1);
323 hold on
324     plot(tid, alphas, 'LineWidth',2,'Color','black');
325     plot(tid, alpCI, '—r', 'LineWidth',2);
326 hold off
327 title('\alpha_t');
328 subplot(1,2,2);
329 hold on
330     plot(tid, tauhat, 'LineWidth',2,'Color','black');
331     plot(tid, tauCI, '—r', 'LineWidth',2);
332 hold off
333 title('\tau_t');
334
335 apt_rate = [counth countlam]/nloop;
336 clc;
337 fprintf('\n');

```

```

338 fprintf('Parámetro          | Media a posteriori | Intervalo de confianza al 95 %%\n');
339 fprintf('mu                | %.3f              | (%.3f, %.3f)\n', thetahat(1), ...
340         thetaCI(1,1), thetaCI(1,2));
341 fprintf('beta              | %.3f              | (%.3f, %.3f)\n', thetahat(2), ...
342         thetaCI(2,1), thetaCI(2,2));
343 fprintf('phi               | %.3f              | (%.3f, %.3f)\n', thetahat(3), ...
344         thetaCI(3,1), thetaCI(3,2));
345 fprintf('sigma2            | %.3f              | (%.3f, %.3f)\n', thetahat(4), ...
346         thetaCI(4,1), thetaCI(4,2));
347 fprintf('omega2_alpha       | %.3f              | (%.3f, %.3f)\n', thetahat(5), ...
348         thetaCI(5,1), thetaCI(5,2));
349 fprintf('omega_{alpha,tau} | %.3f              | (%.3f, %.3f)\n', thetahat(6), ...
350         thetaCI(6,1), thetaCI(6,2));
351 fprintf('omega2_tau         | %.3f              | (%.3f, %.3f)\n', thetahat(7), ...
352         thetaCI(7,1), thetaCI(7,2));
353
354
355 disp('Cálculo de la verosimilitud marginal.... ');
356 llike0 = fdp_y_dado_h_gamma_psi(y, thetahat, ht, phi0, Vphi, mu0, Vmu, beta0, Vbeta, Vgam, nuOmega, ...
357         nuh, Sh, nloop, burnin);
358 %% log verosimilitud mu and phi
359 llike1 = fdp_mu_phi_dado_h_gamma_beta_sig2_Omega(y, thetahat, ht, phi0, Vphi, mu0, Vmu, beta0, ...
360         Vbeta, Vgam, nuOmega, nuh, Sh, SOmega, nloop, burnin);
361 %% log verosimilitud sigma2 Omega
362 llike2 = fdp_sig2_Omega_dado_h_gamma_mu_phi_beta(y, thetahat, ht, phi0, Vphi, mu0, Vmu, beta0, ...
363         Vbeta, Vgam, nuOmega, nuh, Sh, SOmega, nloop, burnin);
364 %% log verosimilitud beta
365 llike3 = fdp_beta_dado_h_gamma_mu_phi_sig2_Omega(y, thetahat, ht, phi0, Vphi, mu0, Vmu, beta0, ...
366         Vbeta, Vgam, nuOmega, nuh, Sh, SOmega, nloop, burnin);
367 llike=prior(thetahat(1), thetahat(3), thetahat(2), thetahat(4), [thetahat(5), thetahat(6); ...
368 thetahat(6), thetahat(7)])+mean(llike0-llike1-llike2-llike3);
369 fprintf('\n');
370 fprintf('log de la verosimilitud integral: %d\n', mean(llike0));
371 fprintf('\n');
372 fprintf('log de la verosimilitud marginal: %d\n', llike);
373 fprintf('\n');
374 disp( ['MCMC toma ' num2str( etime( clock, start_time ) ) ' segundos' ] );

```

```

1
2 %% Calculo de las probabilidades dinámicas que alpha_t no igual a 0.
3
4 postalp = zeros(T,1);
5 prialp = zeros(T,1);
6
7 Omega = zeros(2,2);
8 disp('Calculo de las probabilidades dinámicas que alpha_t no igual a 0.... ');
9 %% Evaluar la densidad a posteriori en alpha_t igual a 0
10 for loop = 1:nloop - burnin
11     Omega([1 2 4]) = store_theta(loop,5:7);
12     Omega(3) = store_theta(loop,6);
13     h = store_h(loop,:);
14
15     invOmegagam = [invVgam sparse(2,2*(T-1)); ...
16                 sparse(2*(T-1),2) kron(speye(T-1),Omega\speye(2))];
17     Xgam = SURform([exp(h) ones(T,1)]);
18     tmp1 = Xgam' * sparse(1:T,1:T,1./exp(h));
19     Kgam = H'*invOmegagam*H + tmp1*Xgam;
20     Dgam = diag(Kgam\speye(2*T));
21     gamhat = Kgam\(tmp1*y);
22     postalp = postalp + normpdf(0, gamhat(1:2:end), sqrt(Dgam(1:2:end)));
23 end
24
25
26 %% Evaluar la densidad a priori en alpha_t igual a 0
27 N = 5000;
28 H = speye(T) - sparse(2:T,1:(T-1),ones(1,T-1),T,T);
29 for ii=1:N
30     Omega = iwishrnd(SOmega, nuOmega);
31     omega2alp = Omega(1);
32     invOmegaalp = sparse(1:T,1:T,[1/Vgam(1) 1/omega2alp*ones(1,T-1)]);
33     Kalp = H'*invOmegaalp*H;
34     Dalp = diag(Kalp\speye(T));

```



```

35     prialp = prialp + normpdf(0,0,sqrt(Dalp));
36 end
37 BF = postalp/(nloop-burnin) ./ (prialp/N);
38 prob = 1./(1+BF);

1 %% Simulación de la fdp p(mu,phi|h,gamma,beta,sig2,Omega) usando muestreo de Gibbs
2
3 function llike=fdp_mu_phi_dado_h_gamma_beta_sig2_Omega(y,thetahat,ht,phi0,Vphi,mu0,...
4     Vmu,beta0,Vbeta,Vgam,nuOmega,nuh,Sh,SOmega,M,burnin)
5     invVgam = Vgam\speye(2);
6     mu=thetahat(1);
7     beta=thetahat(2);
8     phi=thetahat(3);
9     sig2=thetahat(4);
10    Omega=[thetahat(5),thetahat(6);thetahat(6),thetahat(7)];
11
12    %% compute a few things outside the loop
13    T = length(y);
14    D= repmat(1/sig2,T,1);
15    D(1)=(1-phi^2)/sig2;
16    ish=diag(D);
17
18    H = speye(2*T) - sparse(3:2*T,1:2*(T-1),ones(1,2*(T-1)),2*T,2*T);
19    Hphi = speye(T) - sparse(2:T,1:(T-1),phi*ones(1,T-1),T,T);
20    HinvSH = Hphi'*ish*Hphi;
21    deltah = Hphi\([mu; mu*(1-phi)*ones(T-1,1)] + [0;y(1:end-1)]*beta);
22    HinvSHdeltah = HinvSH*deltah;
23
24    newnuh = T/2 + nuh;
25    counth = 0; countlam = 0;
26    h = mu + sqrt(sig2)*randn(T,1);
27    expht = exp(h);
28    llike=zeros(M,1);
29 for i = 1:M
30     %% sample gam
31     invOmegagam = [invVgam sparse(2,2*(T-1)); ...
32         sparse(2*(T-1),2) kron(speye(T-1),Omega\speye(2))];
33     Xgam = SURform([expht ones(T,1)]);
34     temp1 = Xgam' * sparse(1:T,1:T,1./expht);
35     Kgam = H' * invOmegagam * H + temp1 * Xgam;
36     Cgam = chol(real(Kgam),'lower');
37     gamhat = Kgam\((temp1*y));
38     gam = gamhat + Cgam'\randn(2*T,1);
39     alp = gam(1:2:end);
40     tau = gam(2:2:end);
41
42     %% sample h
43     D= repmat(1/sig2,T,1);
44     D(1)=(1-phi^2)/sig2;
45     ish=diag(D);
46     HinvSH = Hphi'*ish*Hphi;
47     deltah = Hphi\([mu; mu*(1-phi)*ones(T-1,1)] + [0;y(1:end-1)]*beta);
48     HinvSHdeltah = HinvSH*deltah;
49     s2 = (y-tau).^2;
50     errh = 1; ht = h;
51     while errh > 0.001;
52         expht = exp(ht);
53         sinvexpht = s2./expht;
54         alp2expht = alp.^2.*expht;
55         fh = -.5 + .5*sinvexpht - .5*alp2expht;
56         Gh = .5*sinvexpht + .5*alp2expht;
57         Kh = HinvSH + spdiags(Gh,0,T,T);
58         newht = Kh\(fh+Gh.*ht+HinvSHdeltah);
59         errh = norm((newht-h));
60         ht = newht;
61     end
62     cholHh = chol(real(Kh),'lower');
63     %% AR-step:
64     hstar = ht;
65     uh = hstar-deltah;
66     logc = -.5*uh'*HinvSH*uh -.5*sum(hstar) + ...
67         -.5*exp(-hstar)'*(y-tau-alp.*exp(hstar)).^2 + log(3);

```

```

68 logc=real(logc);
69 flag = 0;
70 while flag == 0
71     hc = ht + cholHh'\randn(T,1);
72     vhc = hc-ht;
73     uhc = hc-deltah;
74     alpARc = -.5*uhc'*HinvSH*uhc -.5*sum(hc) + ...
75             -.5*exp(-hc)*(y-tau-alp.*exp(hc)).^2 + .5*vhc'*Kh*vhc - logc;
76     if alpARc > log(rand)
77         flag = 1;
78     end
79 end
80 %MH-step
81 vh = hc-ht;
82 uh = hc-deltah;
83 alpAR = -.5*uh'*HinvSH*uh -.5*sum(hc) + ...
84         -.5*exp(-hc)*(y-tau-alp.*exp(hc)).^2 + .5*vh'*Kh*vh - logc;
85 if alpAR < 0
86     alpMH = 1;
87 elseif alpARc < 0
88     alpMH = - alpAR;
89 else
90     alpMH = alpARc - alpAR;
91 end
92 if alpMH > log(rand) || loop == 1
93     h = hc;
94     expH = exp(hc);
95     counth = counth + 1;
96 end
97
98 %% sample beta
99 ybeta = h(2:end) - mu - phi*(h(1:end-1)-mu);
100 Dbeta = real(1/(1/Vbeta + y(1:end-1)*y(1:end-1)/sig2));
101 betahat = Dbeta*(beta0/Vbeta + y(1:end-1)*ybeta/sig2);
102 beta = betahat + sqrt(Dbeta)*randn;
103
104 %% sample Omega
105 err = reshape(gam(3:end, :) - gam(1:end-2, :), 2, T-1)';
106 Omega = real(iwishrnd(SOmega + err'*err, nuOmega + T-1));
107
108 %% sample sig2
109 errh = [(h(1)-mu)*sqrt(1-phi^2); h(2:end)-phi*h(1:end-1)-mu*(1-phi)-y(1:end-1)*beta];
110 newSh = Sh + sum(errh.^2)/2;
111 sig2 = 1/gamrnd(newnuh, 1./newSh);
112
113 %% sample mu and phi jointly
114 sum1 = sum(h(2:end));
115 sum2 = sum(h(1:end-1));
116 flam = @(x) -(1-x(2))^2/(2*sig2)*(h(1)-x(1))^2 - 1/(2*sig2) * ...
117         sum((h(2:end) - x(2)*h(1:end-1) - x(1)*(1-x(2))-y(1:end-1)*beta).^2) ...
118         - 1/(2*Vphi)*(x(2)-phi0)^2 - 1/(2*Vmu)*(x(1)-mu0)^2;
119 [lamc, g] = proplam(h, sig2, phi0, Vphi, mu0);
120 MHprob = flam(lamc) - flam([mu; phi]) + g([mu; phi]) - g(lamc);
121 if exp(MHprob) > rand
122     mu = lamc(1);
123     phi = lamc(2);
124     countlam = countlam+1;
125 end
126 Hphi = speye(T) - sparse(2:T, 1:(T-1), phi*ones(1, T-1), T, T);
127
128 llike(i) = -0.5*T*log(2*pi*sig2) - (1-phi)^2/(2*sig2)*(h(1)-thetahat(1))^2 ...
129         - sum((h(2:end)-thetahat(3)*h(1:end-1)-thetahat(1)*(1-thetahat(3))) ...
130         - beta*y(1:end-1)).^2)/(2*sig2) - 0.5*log(2*pi*Vphi) - 0.5 ...
131         *(thetahat(3)-phi0).^2 - 0.5*log(2*pi*Vmu) - 0.5*(thetahat(1)-mu0).^2/Vmu;
132
133
134 end
135 llike = (llike((burnin+1):end));
136
137 end

```

1 %% Simulación de la fdp $p(\beta | y, h, \gamma, \mu, \phi, \sigma^2, \Omega)$ usando muestreo de Gibbs

```

2
3 function llike=fdp_beta_dado_h_gamma_mu_phi_sig2_Omega(y, thetatah , ht , phi0 , Vphi , mu0 , Vmu , ...
4     beta0 , Vbeta , Vgam , nuOmega , nuh , Sh , SOmega , M , burnin )
5     invVgam = Vgam \ speye(2) ;
6     mu=thetatah(1) ;
7     beta=thetatah(2) ;
8     phi=thetatah(3) ;
9     sig2=thetatah(4) ;
10    Omega=[ thetatah(5) , thetatah(6) ; thetatah(6) , thetatah(7) ] ;
11
12    %% compute a few things outside the loop
13    T = length(y) ;
14    D=repmat(1/sig2 , T,1) ;
15    D(1)=(1-phi^2)/sig2 ;
16    ish=diag(D) ;
17
18    H = speye(2*T) - sparse(3:2*T,1:2*(T-1), ones(1,2*(T-1)),2*T,2*T) ;
19    Hphi = speye(T) - sparse(2:T,1:(T-1), phi*ones(1,T-1),T,T) ;
20    HinvSH = Hphi' * ish * Hphi ;
21    deltah = Hphi \ ([mu; mu*(1-phi)*ones(T-1,1)] + [0;y(1:end-1)] * beta) ;
22    HinvSHdeltah = HinvSH * deltah ;
23
24    newnuh = T/2 + nuh ;
25    counth = 0 ; countlam = 0 ;
26    h = mu + sqrt(sig2) * randn(T,1) ;
27    expht = exp(h) ;
28    llike=zeros(M,1) ;
29 for i = 1:M
30     %% sample gam
31     invOmegagam = [invVgam sparse(2,2*(T-1)); ...
32         sparse(2*(T-1),2) kron(speye(T-1),Omega\speye(2))];
33     Xgam = SURform([expht ones(T,1)]);
34     temp1 = Xgam' * sparse(1:T,1:T,1./expht);
35     Kgam = H' * invOmegagam * H + temp1 * Xgam;
36     Cgam = chol(real(Kgam), 'lower');
37     gamhat = Kgam \ (temp1 * y);
38     gam = gamhat + Cgam' \ randn(2*T,1);
39     alp = gam(1:2:end);
40     tau = gam(2:2:end);
41
42     %% sample h
43     D=repmat(1/sig2 , T,1) ;
44     D(1)=(1-phi^2)/sig2 ;
45     ish=diag(D) ;
46     HinvSH = Hphi' * ish * Hphi ;
47     deltah = Hphi \ ([mu; mu*(1-phi)*ones(T-1,1)] + [0;y(1:end-1)] * beta) ;
48     HinvSHdeltah = HinvSH * deltah ;
49     s2 = (y-tau).^2 ;
50     errh = 1 ; ht = h ;
51     while errh > 0.001 ;
52         expht = exp(ht) ;
53         sinvexpht = s2 ./ expht ;
54         alp2expht = alp.^2 .* expht ;
55         fh = -.5 + .5 * sinvexpht - .5 * alp2expht ;
56         Gh = .5 * sinvexpht + .5 * alp2expht ;
57         Kh = HinvSH + spdiags(Gh,0,T,T) ; %% spdiags pone diagonal a Gh q es vector
58         newht = Kh \ (fh + Gh .* ht + HinvSHdeltah) ;
59         errh = norm((newht-ht)) ;
60         ht = newht ;
61     end
62     cholHh = chol(real(Kh), 'lower') ;
63     %% AR-step :
64     hstar = ht ;
65     uh = hstar - deltah ;
66     logc = -.5 * uh' * HinvSH * uh - .5 * sum(hstar) + ...
67         - .5 * exp(-hstar) * (y-tau - alp .* exp(hstar)).^2 + log(3) ;
68     logc = real(logc) ;
69     flag = 0 ;
70     while flag == 0
71         hc = ht + cholHh' \ randn(T,1) ;
72         vhc = hc - ht ;
73         uhc = hc - deltah ;

```

```

74     alpARc = -.5*uhc'*HinvSH*uhc -.5*sum(hc) + ...
75             -.5*exp(-hc)'*(y-tau-alp.*exp(hc)).^2 + .5*vhc'*Kh*vhc - logc;
76     if alpARc > log(rand)
77         flag = 1;
78     end
79 end
80 %%MH-step
81 vh = hc-ht;
82 uh = hc-deltah;
83 alpAR = -.5*uh'*HinvSH*uh -.5*sum(hc) + ...
84         -.5*exp(-hc)'*(y-tau-alp.*exp(hc)).^2 + .5*vh'*Kh*vh - logc;
85 if alpAR < 0
86     alpMH = 1;
87 elseif alpARc < 0
88     alpMH = - alpAR;
89 else
90     alpMH = alpARc - alpAR;
91 end
92 if alpMH > log(rand) || loop == 1
93     h = hc;
94     expH = exp(hc);
95     counth = counth + 1;
96 end
97
98 %% sample beta
99 ybeta = h(2:end) - mu - phi*(h(1:end)-mu);
100 Dbeta = real(1/(1/Vbeta + y(1:end-1)*y(1:end-1)/sig2));
101 betahat = Dbeta*(beta0/Vbeta + y(1:end-1)*ybeta/sig2);
102 beta = betahat + sqrt(Dbeta)*randn;
103
104 %% Puntos de alta densidad
105 mu=thetahat(1);
106 beta=thetahat(2);
107 phi=thetahat(3);
108 sig2=thetahat(4);
109 Omega=[thetahat(5),thetahat(6);thetahat(6),thetahat(7)];
110 llike(i)=-log(2*pi*Dbeta)/2-(beta-betahat)^2/(2*Dbeta);
111
112 end
113 llike=(llike((burnin+1):end));
114
115 end

```

```

1 %% Simulación de la fdp p(sig2*,Omega*|y,h,gamma,beta,mu,phi) usando muestreo de Gibbs
2
3 function llike=fdp_sig2_Omega_dado_h_gamma_mu_phi_beta(y,thetahat,ht,phi0,Vphi,mu0,Vmu,...
4             beta0,Vbeta,Vgam,nuOmega,nuh,Sh,SOMega,M,burnin)
5     %% initialize the Markov chain
6     invVgam = Vgam\speye(2);
7     mu=thetahat(1);
8     beta=thetahat(2);
9     phi=thetahat(3);
10    sig2=thetahat(4);
11    Omega=[thetahat(5),thetahat(6);thetahat(6),thetahat(7)];
12
13    %% compute a few things outside the loop
14    T = length(y);
15    D= repmat(1/sig2,T,1);
16    D(1)=(1-phi^2)/sig2;
17    ish=diag(D);
18    h = mu + sqrt(sig2)*randn(T,1);
19    expH = exp(h);
20    H = speye(2*T) - sparse(3:2*T,1:2*(T-1),ones(1,2*(T-1)),2*T,2*T);
21    Hphi = speye(T) - sparse(2:T,1:(T-1),phi*ones(1,T-1),T,T);
22    HinvSH = Hphi'*ish*Hphi;
23    deltah = Hphi\([mu; mu*(1-phi)*ones(T-1,1)] + [0;y(1:end-1)]*beta);
24    HinvSHdeltah = HinvSH*deltah;
25    newnuh = T/2 + nuh;
26    counth = 0; countlam = 0;
27    llike=zeros(M,1);
28 for i = 1:M
29     %% sample gam

```

```

30 invOmegagam = [invVgam sparse(2,2*(T-1)); ...
31     sparse(2*(T-1),2) kron(speye(T-1),Omega\speye(2))];
32 Xgam = SURform([exph ones(T,1)]);
33 temp1 = Xgam' * sparse(1:T,1:T,1./exph);
34 Kgam = H'*invOmegagam*H + temp1*Xgam;
35 Cgam = chol(real(Kgam),'lower');
36 gamhat = Kgam\(temp1*y);
37 gam = gamhat + Cgam'\randn(2*T,1);
38 alp = gam(1:2:end);
39 tau = gam(2:2:end);
40
41 %% sample h
42 D= repmat(1/sig2,T,1);
43 D(1)=(1-phi^2)/sig2;
44 ish=diag(D);
45 HinvSH = Hphi'*ish*Hphi;
46 deltah = Hphi\([mu; mu*(1-phi)*ones(T-1,1)] + [0;y(1:end-1)]*beta);
47 HinvSHdeltah = HinvSH*deltah;
48 s2 = (y-tau).^2;
49 errh = 1; ht = h;
50 while errh > 0.001;
51     expht = exp(ht);
52     sinvexpht = s2./expht;
53     alp2expht = alp.^2.*expht;
54     fh = -.5 + .5*sinvexpht - .5*alp2expht;
55     Gh = .5*sinvexpht + .5*alp2expht;
56     Kh = HinvSH + spdiags(Gh,0,T,T); %% spdiags pone diagonal a Gh q es vector
57     newht = Kh\(fh+Gh.*ht+HinvSHdeltah);
58     errh = norm((newht-h));
59     ht = newht;
60 end
61 cholHh = chol(real(Kh),'lower');
62 %% AR-step:
63 hstar = ht;
64 uh = hstar-deltah;
65 logc = -.5*uh'*HinvSH*uh -.5*sum(hstar) + ...
66     -.5*exp(-hstar)*(y-tau-alp.*exp(hstar)).^2 + log(3);
67 logc=real(logc);
68 flag = 0;
69 while flag == 0
70     hc = ht + cholHh'\randn(T,1);
71     vhc = hc-ht;
72     uhc = hc-deltah;
73     alpARc = -.5*uhc'*HinvSH*uhc -.5*sum(hc) + ...
74     -.5*exp(-hc)*(y-tau-alp.*exp(hc)).^2 + .5*vhc'*Kh*vhc - logc;
75     if alpARc > log(rand)
76         flag = 1;
77     end
78 end
79 %% MH-step
80 vh = hc-ht;
81 uh = hc-deltah;
82 alpAR = -.5*uh'*HinvSH*uh -.5*sum(hc) + ...
83     -.5*exp(-hc)*(y-tau-alp.*exp(hc)).^2 + .5*vh'*Kh*vh - logc;
84 if alpAR < 0
85     alpMH = 1;
86 elseif alpARc < 0
87     alpMH = -alpAR;
88 else
89     alpMH = alpARc - alpAR;
90 end
91 if alpMH > log(rand) || loop == 1
92     h = hc;
93     exph = exp(hc);
94     counth = counth + 1;
95 end
96
97 %% sample beta
98 ybeta = h(2:end) - mu - phi*(h(1:end-1)-mu);
99 Dbeta = real(1/(1/Vbeta + y(1:end-1)*y(1:end-1)/sig2));
100 betahat = Dbeta*(beta0/Vbeta + y(1:end-1)*ybeta/sig2);
101 beta = betahat + sqrt(Dbeta)*randn;

```

```

102
103
104 %% Puntos de alta densidad
105 mu=thetahat(1);
106 phi=thetahat(3);
107 sig2=thetahat(4);
108 Omega=[thetahat(5),thetahat(6);thetahat(6),thetahat(7)];
109
110 err = reshape(gam(3:end,:) - gam(1:end-2,:),2,T-1)';
111 errh = [(h(1)-mu)*sqrt(1-phi^2); h(2:end)-phi*h(1:end-1)-mu*(1-phi)-y(1:end-1)*beta];
112 newSh = Sh + sum(errh.^2)/2; %% S_sig2 techo
113 SOmegahat=SOmega+ err' * err; %% S_Omega
114 v=nuOmega+T-1;
115 a1=(-(v+2+1)/2)*log(det(SOmegahat/(v+2+1)));
116 a2=-0.5*trace(SOmegahat*inv(SOmegahat/(v+2+1)));
117 a3=(v/2)*log(det(SOmegahat))-log(2^(v))-log(pi^(0.5))-gammaln(v/2)-gammaln(v/2-0.5);
118 b=newnuh*log(newSh)+(-newnuh-1)*log(sig2)-(newSh)/sig2-log(gamma(newnuh));
119 llike(i)=a1+a2+a3+b;
120
121 %% sample Omega
122 err = reshape(gam(3:end,:) - gam(1:end-2,:),2,T-1)';
123 Omega =real(iwishrnd(SOmega + err' * err,nuOmega + T-1));
124
125 %% sample sig2
126 errh = [(h(1)-mu)*sqrt(1-phi^2); h(2:end)-phi*h(1:end-1)-mu*(1-phi)-y(1:end-1)*beta];
127 newSh = Sh + sum(errh.^2)/2;
128 sig2 = 1/gamrnd(newnuh, 1./newSh);
129 end
130 llike=(llike((burnin+1):end));
131
132 end

```

```

1 %% Simulación de la fdp p(y|h,gamma,psi*) usando muestreo de Gibbs
2
3 function llike = fdp_y_dado_h_gamma_psi(y,thetahat,ht,phi0,Vphi,mu0,Vmu,beta0,...
4 Vbeta,Vgam,nuOmega,nuh,Sh,M,burnin)
5 invVgam = Vgam\speye(2);
6 mu=thetahat(1);
7 beta=thetahat(2);
8 phi=thetahat(3);
9 sig2=thetahat(4);
10 Omega=[thetahat(5),thetahat(6);thetahat(6),thetahat(7)];
11
12
13 %% compute a few things outside the loop
14 T = length(y);
15 D=repmat(1/sig2,T,1);
16 D(1)=(1-phi^2)/sig2;
17 ish=diag(D);
18
19 H = speye(2*T) - sparse(3:2*T,1:2*(T-1),ones(1,2*(T-1)),2*T,2*T);
20 Hphi = speye(T) - sparse(2:T,1:(T-1),phi*ones(1,T-1),T,T);
21 HinvsH = Hphi' * ish * Hphi;
22 deltah = Hphi \ ([mu; mu*(1-phi)*ones(T-1,1)] + [0;y(1:end-1)] * beta);
23 HinvsHdeltah = HinvsH * deltah;
24
25
26 newnuh = T/2 + nuh;
27 counth = 0; countlam = 0;
28 h = mu + sqrt(sig2)*randn(T,1);
29 exp_h = exp(h);
30 llike=zeros(M,1);
31 for i = 1:M
32 %% sample gam
33 invOmegagam = [invVgam sparse(2,2*(T-1)); ...
34 sparse(2*(T-1),2) kron(speye(T-1),Omega\speye(2))];
35 Xgam = SURform([exp_h ones(T,1)]);
36 temp1 = Xgam' * sparse(1:T,1:T,1./exp_h);
37 Kgam = H' * invOmegagam * H + temp1 * Xgam;
38 Cgam = chol(real(Kgam),'lower');
39 gamhat = Kgam \ (temp1 * y);
40 gam = gamhat + Cgam' \ randn(2*T,1);

```

```

41 alp = gam(1:2:end);
42 tau = gam(2:2:end);
43
44 %% sample h
45 D= repmat(1/sig2 ,T,1);
46 D(1)=(1-phi^2)/sig2;
47 ish=diag(D);
48 HinvSH = Hphi' * ish * Hphi;
49 deltah = Hphi \ ([mu; mu*(1-phi)*ones(T-1,1)] + [0;y(1:end-1)] * beta);
50 HinvSHdeltah = HinvSH * deltah;
51 s2 = (y-tau).^2;
52 errh = 1; ht = h;
53 while errh > 0.001;
54     expht = exp(ht);
55     sinvexpht = s2 ./ expht;
56     alp2expht = alp.^2 * expht;
57     fh = -.5 + .5 * sinvexpht - .5 * alp2expht;
58     Gh = .5 * sinvexpht + .5 * alp2expht;
59     Kh = HinvSH + spdiags(Gh,0,T,T); %% spdiags pone diagonal a Gh q es vector
60     newht = Kh \ (fh + Gh .* ht + HinvSHdeltah);
61     errh = norm((newht-ht));
62     ht = newht;
63 end
64 cholHh = chol(real(Kh), 'lower');
65 %% AR-step:
66 hstar = ht;
67 uh = hstar - deltah;
68 logc = -.5 * uh' * HinvSH * uh - .5 * sum(hstar) + ...
69     - .5 * exp(-hstar) * (y-tau - alp .* exp(hstar)).^2 + log(3);
70 logc = real(logc);
71 flag = 0;
72 while flag == 0
73     hc = ht + cholHh' \ randn(T,1);
74     vhc = hc - ht;
75     uhc = hc - deltah;
76     alpARc = -.5 * uhc' * HinvSH * uhc - .5 * sum(hc) + ...
77         - .5 * exp(-hc) * (y-tau - alp .* exp(hc)).^2 + .5 * vhc' * Kh * vhc - logc;
78     if alpARc > log(rand)
79         flag = 1;
80     end
81 end
82 %% MH-step
83 vh = hc - ht;
84 uh = hc - deltah;
85 alpAR = -.5 * uh' * HinvSH * uh - .5 * sum(hc) + ...
86     - .5 * exp(-hc) * (y-tau - alp .* exp(hc)).^2 + .5 * vh' * Kh * vh - logc;
87 if alpAR < 0
88     alpMH = 1;
89 elseif alpARc < 0
90     alpMH = -alpAR;
91 else
92     alpMH = alpARc - alpAR;
93 end
94 if alpMH > log(rand) || loop == 1
95     h = hc;
96     expht = exp(hc);
97     counth = counth + 1;
98 end
99
100
101 s2 = (y-tau).^2;
102 expht = exp(ht);
103 sinvexpht = s2 ./ expht;
104 alp2expht = alp.^2 * expht;
105 fh = -.5 + .5 * sinvexpht - .5 * alp2expht;
106 Gh = .5 * sinvexpht + .5 * alp2expht;
107 llike(i) = -.5 * T * log(2 * pi) - 0.5 * sum(ht) - 0.5 * (exp(-ht) .* (y-tau - alp .* exp(ht)))' * ...
108     (y-tau - alp .* exp(ht)) + (hc-ht)' * fh - 0.5 * (hc-ht)' * (Gh .* (hc-ht));
109
110 end
111 llike = (llike((burnin+1):end));
112 end

```

```

1
2 %Esta función obtiene una desidad candidata para muestrear (mu, phi) usando
3 %un paso MH
4 %Detalles: se utiliza el método de Newton–Raphason (con BHHH) para maximizar
5 %la distribución condicional parmetrizada de (mu, phi) como
6 %delta = (mu, tanh^(−1)(phi))
7 %para obtener las muestras de (mu, phi) de regreso se utiliza el teorema de
8 %cambio de variable
9 %Salida: una muestra extraída de (mu, phi) y la densidad candidata propuesta
10
11 function [lam g] = proplam(h, sig2, phi0, Vphi, mu0)
12
13 maxcount = 100;
14 T = size(h,1);
15 s = zeros(T,2);
16 mut = mean(h);
17 phit = (h(1:end−1)−mut)'*(h(2:end)−mut)/sum((h(1:end−1)−mut).^2);
18 delt = [mut atanh(phit)]';
19 count = 0;
20 flag = 0;
21 while flag == 0
22     s(1,1) = sech(delt(2))^2/sig2*(h(1)−delt(1))−(delt(1)−mu0)/Vphi;
23     s(2:T,1) = (1−tanh(delt(2)))/sig2 * ...
24         (h(2:end)−tanh(delt(2))*h(1:end−1) − delt(1)*(1−tanh(delt(2))));
25     s(1,2) = sech(delt(2))^2*tanh(delt(2))*((h(1)−delt(1))^2/sig2−1/Vphi) ...
26         +phi0/Vphi*sech(delt(2))^2;
27     s(2:T,2) = sech(delt(2))^2/sig2 * (h(1:end−1)−delt(1)) .* ...
28         (h(2:T)−delt(1)−tanh(delt(2))*(h(1:end−1)−delt(1)));
29     S = sum(s)';
30     B = s'*s;
31     delt = delt + B\S;
32     count = count + 1;
33     if sum(abs(B\S)) > 10^(−4) && count < maxcount %stopping criteria
34         flag = 1;
35     end
36 end
37 [C,p] = chol(B,'lower');
38 if count == maxcount || p~=0
39     delt = [mut atanh(phit)]';
40     s(1,1) = sech(delt(2))^2/sig2*(h(1)−delt(1));
41     s(2:T,1) = (1−tanh(delt(2)))/sig2 * ...
42         (h(2:end)−tanh(delt(2))*h(1:end−1) − delt(1)*(1−tanh(delt(2))));
43     s(1,2) = sech(delt(2))^2*tanh(delt(2))/sig2 * (h(1)−delt(1));
44     s(2:T,2) = sech(delt(2))^2/sig2 * (h(1:end−1)−delt(1)) .* ...
45         (h(2:T)−delt(1)−tanh(delt(2))*(h(1:end−1)−delt(1)));
46     B = s'*s;
47     C = chol(B,'lower');
48 end
49 %t proposal with df nu = 5
50 nu = 5;
51 del = delt + C*\randn(2,1)/sqrt(gamrnd(nu/2,2/nu));
52 lam = [del(1); tanh(del(2))];
53 c = gammaln((nu+2)/2) − gammaln(nu/2) − log(nu) − log(pi) + .5*log(det(B));
54
55 g = @(x) c + 2*log(cosh(atanh(x(2)))) ... %Jacobian of transformation
56 − (nu+2)/2 * log( 1 + 1/nu * ([x(1) atanh(x(2))]−delt)')*inv(B)*([x(1); atanh(x(2))]−delt));
57
58 end

```

```

1
2 %Esta función construye la matriz de regresión X, así que la regresión
3 %está en la forma SUR %
4 %mirar (Chan, J.C.C. (2017))
5 function Xout = SURform(X)
6 [r c] = size(X);
7
8 idi = kron((1:r)',ones(c,1));
9 idj = (1:r*c)';
10
11 Xout = sparse(idi, idj, reshape(X', r*c, 1));
12
13 end

```




ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE CIENCIAS

CARRERA DE INGENIERÍA MATEMÁTICA

ORDEN DE EMPASTADO

De acuerdo a lo estipulado en el artículo 27 del Instructivo para la Implementación de la Unidad de Titulación en las Carreras y Programas vigentes de la Escuela Politécnica Nacional, aprobado por Consejo de Docencia en sesión extraordinaria del 29 de abril de 2015 y una vez verificado el cumplimiento del formato de presentación establecido, se autoriza la impresión y encuadernación final del Trabajo de Titulación presentado por el señor **PABLO DAVID ALBÁN CERDA**.

Fecha de autorización: Quito, D.M., 24 de julio de 2018.


Dr. Luis Lascano
DECANO
FACULTAD DE CIENCIAS

