



UNIVERSIDAD NACIONAL DE COLOMBIA

Una Aplicación del Método de Momentos Eficiente a la Estimación de Modelos de Volatilidad Estocástica

Édison Hernán Ruiz Osorno

Universidad Nacional de Colombia

Facultad de Ciencias

Escuela de Estadística

Medellín, Colombia

2012

Una Aplicación del Método de Momentos Eficiente a la Estimación de Modelos de Volatilidad Estocástica

Por:

Édison Hernán Ruiz Osorno
ehruizo@unal.edu.co

Director:

Norman Diego Giraldo, M.Sc.

Trabajo de grado presentado como requisito parcial para optar al título de
Magíster en Ciencias-Estadística

Línea de investigación:
Econometría Financiera

Universidad Nacional de Colombia
Facultad de Ciencias
Escuela de Estadística
Medellín, Colombia
2012

©2012 Édison H. Ruiz

Agradecimientos

Quiero agradecer a quienes aportaron profesional y personalmente durante estos casi dos años y medio de formación académica. En primer lugar, al director de la tesis, profesor Norman Diego Giraldo, por su excelente orientación y sus valiosas ideas. A los jurados, Mauricio Lopera Castaño de la Universidad de Antioquia y Elkin Castaño Vélez de la Universidad Nacional de Colombia, sede Medellín, por su revisión rigurosa y sus acertados comentarios sobre una versión preliminar de esta tesis. También, en un sentido más personal, agradezco a mis compañeros de la Maestría en Ciencias-Estadística, con quienes tuve la oportunidad de compartir durante estos meses. Finalmente, a los demás profesores y personal administrativo de la Escuela de Estadística, quienes de una u otra forma facilitaron mi crecimiento académico y personal.

Medellín, Colombia
Mayo de 2012

Édison H. Ruiz

Resumen

Los modelos de volatilidad estocástica tratan de explicar las características de las series de retornos de acciones y otros activos financieros. En estos modelos, la volatilidad se considera un proceso aleatorio latente, el cual determina, en parte, las propiedades de la distribución de los retornos. Pese a que en teoría reproducen bien las principales características de los datos, los modelos de volatilidad estocástica tienen la dificultad que no permiten deducir una expresión analítica para la función de verosimilitud, por lo que su estimación se debe abordar desde otros enfoques. Uno de los métodos que ha sido propuesto es el Método de Momentos Eficiente (EMM), cuya idea fundamental es utilizar los valores esperados de los scores de una densidad auxiliar bajo el modelo de volatilidad estocástica como condiciones de momentos en un criterio GMM. En la medida que la densidad auxiliar se ajusta mejor a los datos, el estimador EMM es más eficiente. El objetivo de esta tesis es llevar a cabo un estudio Monte Carlo de la capacidad del EMM para estimar modelos de volatilidad estocástica con respuesta asimétrica de la volatilidad a cambios en los precios (efecto *leverage*). Entre los modelos auxiliares empleados en la estimación, se propone la utilización del modelo ARCH asimétrico (APARCH), debido a que éste puede reproducir el efecto de asimetría, entre otras características observadas en los datos.

Palabras clave: Volatilidad Estocástica, Método de Momentos Eficiente, Monte Carlo, APARCH, EGARCH.

Abstract

Stochastic volatility models try to explain observed features of stock returns and other financial assets. These models treat volatility as a latent stochastic process that partly determines the properties of the distribution of returns. Due to the absence of a likelihood function, estimation of stochastic volatility models must be tackled from alternative approaches. One of the methods that has been proposed is the Efficient Method of Moments (EMM), whose basic idea is to use the expectation of the scores of an auxiliary density under the stochastic volatility model as moment conditions in a GMM criterion. As the auxiliary density better fits the data, the EMM estimator becomes more efficient. The objective of this thesis is to carry out a Monte Carlo study of the capability of EMM to estimate stochastic volatility models with leverage effect. Among the auxiliary models considered, we propose the use of the APARCH model, since it can reproduce the asymmetric effect, among other observed features of the data.

Keywords: Stochastic Volatility, Efficient Method of Moments, Monte Carlo, APARCH, EGARCH.

Índice general

Agradecimientos	v
Resumen	vi
Índice general	viii
Índice de figuras	x
Índice de tablas	xii
1. Introducción	1
1.1. Introducción General	1
1.2. Revisión del Problema de Investigación	6
1.2.1. Comparación de modelos auxiliares	8
1.2.2. Síntesis	18
2. El Modelo de Volatilidad Estocástica	19
2.1. Introducción	19
2.2. Propiedades Estadísticas	21
2.3. Extensiones del Modelo Básico	26

3. El Método de Momentos Eficiente	30
3.1. Introducción	30
3.2. Descripción y Propiedades del EMM	32
3.2.1. Propiedades asintóticas del estimador EMM y pruebas de especificación	35
3.3. Síntesis	36
4. Simulaciones y Estimaciones	37
4.1. Introducción	37
4.2. Configuración de las Simulaciones	38
4.2.1. Scores	42
4.2.2. Procedimiento de estimación	44
4.3. Resultados	46
4.3.1. Modelo SV log-normal	46
4.3.2. Modelo SV con colas pesadas	57
4.4. Síntesis	63
5. Conclusiones	65
Bibliografía	69
A. Resultados de las estimaciones para tamaños de muestra 1.000 y 2.000	76
B. Código en R utilizado para las estimaciones	83

Índice de figuras

1.1. Promedio acumulado de la varianza condicional estimada para el modelo de volatilidad estocástica con varianza condicional “aleatoria”. . .	12
1.2. Promedio acumulado de la varianza condicional estimada para los modelos de volatilidad estocástica con varianza condicional “persistente” y “altamente persistente”.	13
1.3. Inestabilidad de la ecuación recursiva del EGARCH.	14
4.1. Resultados de las estimaciones EMM para el tamaño de muestra 500 del modelo SV log-normal.	51
4.2. Resultados de las estimaciones EMM para el tamaño de muestra 4.000 del modelo SV log-normal.	52
4.3. Comportamiento de los estimadores EMM a medida que aumenta el tamaño de muestra.	53
4.4. QQ-plots normales de las distribuciones de los estimadores EMM para el tamaño de muestra 500 en el modelo SV log-normal.	55
4.5. QQ-plots normales de las distribuciones de los estimadores EMM para el tamaño de muestra 4.000 en el modelo SV log-normal.	56
4.6. Resultados de las estimaciones EMM para el tamaño de muestra 4.000 del modelo SV con colas pesadas.	60
4.7. QQ-plots normales de las distribuciones de los estimadores EMM para el tamaño de muestra 4.000 en el modelo SV con colas pesadas. . . .	62

A.1. Resultados de las estimaciones EMM para el tamaño de muestra 1.000 del modelo SV log-normal.	77
A.2. Resultados de las estimaciones EMM para el tamaño de muestra 2.000 del modelo SV log-normal.	78
A.3. QQ-plots normales de las distribuciones de los estimadores EMM para el tamaño de muestra 1.000 en el modelo SV log-normal.	79
A.4. QQ-plots normales de las distribuciones de los estimadores EMM para el tamaño de muestra 2.000 en el modelo SV log-normal.	80
A.5. Resultados de las estimaciones EMM para el tamaño de muestra 2.000 del modelo SV con colas pesadas.	81
A.6. QQ-plots normales de las distribuciones de los estimadores EMM para el tamaño de muestra 2.000 en el modelo SV con colas pesadas.	82

Índice de tablas

1.1. Parámetros de los modelos de volatilidad estocástica simulados. . . .	11
1.2. Distribución de frecuencias de las desviaciones porcentuales de las volatilidades estimadas con respecto a las verdaderas volatilidades condicionales generadas por los modelos de volatilidad estocástica.	17
4.1. Modelos auxiliares empleados en la estimación por EMM del modelo de volatilidad estocástica.	46
4.2. Resultados de las estimaciones EMM del modelo de volatilidad estocástica log-normal.	49
4.3. Resultados de las estimaciones EMM del modelo de volatilidad estocástica con colas pesadas.	59

Capítulo 1

Introducción

1.1. Introducción General

El modelamiento de los retornos financieros ha generado durante décadas una gran cantidad de investigación. Esto se debe fundamentalmente a que este tema juega un rol central en áreas como el análisis de riesgo, conformación de portafolios, pronóstico de la volatilidad y valoración de opciones, entre otros. Los primeros modelos asumían que los retornos constituyen una secuencia de variables aleatorias independientes e idénticamente distribuidas (iid). Osborne (1959) propuso la distribución normal, la cual empezó a ser rápidamente cuestionada por Mandelbrot (1963) y Fama (1965), quienes encontraron que los datos de retornos diarios tienen muchas observaciones cercanas a un valor central y muchas observaciones extremas que no pueden ser explicadas por el modelo Gaussiano. Esta observación condujo a Mandelbrot (1963) a proponer la familia de distribuciones Pareto estables, la cual puede modelar curtosis alta y colas pesadas, pero tiene el problema que es de varianza infinita, lo que la hace poco atractiva para modelar los retornos.

Más adelante se observó que las autocorrelaciones de los retornos son cercanas a cero (Clark (1973), Taylor (1982, 1986)), pero los retornos al cuadrado y potencias de los retornos absolutos tienen autocorrelaciones significativas que decaen lentamente (Taylor (1986), Ding, Granger y Engle (1993)). Esto quiere decir que el proceso generador de retornos es no lineal, lo que implica descartar la hipótesis de secuencia iid y

sugiere de paso que estos podrían modelarse como un proceso ruido blanco. Por otro lado, también se observó que la varianza de los retornos diarios no es constante, lo que constituye evidencia adicional en contra de la hipótesis iid. No obstante, también se ha detectado que la heteroscedasticidad y las colas pesadas se desvanecen al agregar datos por semanas o meses, en cuyo caso la forma de la distribución se empieza a parecer a la Normal (Cont (2001), Shephard y Andersen (2009)).

Clark (1973) propuso un modelo pionero para explicar la heteroscedasticidad y la distribución leptocúrtica de los retornos. En su modelo, los retornos diarios son la suma de los cambios aleatorios de precios intra-día, los cuales ocurren como resultado de la realización del proceso aleatorio de llegada de información al mercado. En la práctica, este proceso puede estar vinculado con el volumen de comercio y con eventos como el anuncio de dividendos o de datos macroeconómicos, entre otros. En este modelo la curtosis alta se genera por la mezcla de distribuciones, en la que los retornos y la varianza son variables aleatorias. Más allá de si los datos respaldan o no este modelo, su mayor virtud está en que da el soporte teórico necesario a la heteroscedasticidad, al vincularla directamente con el flujo de información.

La principal falencia del modelo de Clark (1973) es que no puede explicar por qué hay periodos de alta y baja volatilidad, fenómeno conocido como agrupamiento de la volatilidad. Este hecho, junto con la evidencia de que los retornos podrían tener las características de un proceso ruido blanco, forman la base para los modelos de heteroscedasticidad condicional (ARCH), propuestos por Engle (1982) y Bollerslev (1986) y el de volatilidad estocástica (SV) propuesto por Taylor (1982, 1986). En los procesos ARCH, la varianza condicional se genera endógenamente como resultado de cambios en los retornos pasados. Por su parte, los modelos SV tratan la varianza condicional como un proceso estocástico no observado, dirigido por una componente aleatoria que reúne los factores exógenos que afectan la varianza. Por ende, en los modelos SV la información relevante para la varianza va más allá de la contenida en los retornos pasados. El agrupamiento de la volatilidad se modela permitiendo que el proceso de varianza esté autocorrelacionado, mientras que la componente exógena es consistente con el modelo de Clark (1973), dado que permite cambios aleatorios en la varianza, los cuales pueden ser generados por el flujo de información.

Un hecho empírico que en los últimos años ha ganado relevancia en la modelación de los retornos es lo que se conoce como efecto de asimetría o *leverage*, y que consiste en que una caída de los retornos produce un aumento mayor de la volatilidad futura

que un aumento de los retornos de la misma magnitud (Black (1976), Nelson (1991)). En otras palabras, este efecto reproduce el hecho de que la volatilidad tiende a aumentar después de malas noticias y a reducirse después de buenas noticias. Numerosos estudios han presentado evidencia del efecto leverage. Engle y Ng (1993) muestran que la relación entre los retornos del día anterior y la varianza actual, relación a la que denominan “curva de impacto de noticias”, es asimétrica, de tal manera que los retornos negativos producen un aumento mayor de la varianza que los retornos positivos. Por su parte, Jacquier, Polson y Rossi (2004) calculan que para el índice CRSP (Center for Research in Security Prices) de Estados Unidos, cuando los retornos caen 2 desviaciones estándar por debajo de la media, la volatilidad es 22 % mayor en los 10 días siguientes que cuando los retornos suben 2 desviaciones estándar por encima de la media. Bollerslev, Litvinova y Tauchen (2006) encuentran que este efecto también se encuentra presente, incluso, en datos de alta frecuencia. Además, hallan que el efecto inverso es insignificante, es decir, que la varianza actual no incide en el nivel de los retornos futuros.

Los modelos que no incorporan el efecto de asimetría, como el GARCH básico, pueden producir malas estimaciones de la varianza para las series en las que este fenómeno está presente. De hecho, Hull y White (1987) argumentan que este efecto es tan significativo que puede explicar en gran medida los sesgos en la valoración de opciones financieras generados por el modelo tradicional de Black y Scholes (1973). Dentro de la familia ARCH, algunos de los modelos que tratan de reproducir el efecto leverage son el GARCH exponencial (EGARCH) de Nelson (1991), el GJR-GARCH de Glosten, Jagannathan y Runkle (1993) y el GARCH asimétrico de potencias (APARCH) de Ding, Granger y Engle (1993). En los modelos de volatilidad estocástica es más sencillo reproducir el efecto leverage, ya que basta con permitir que las innovaciones de las ecuaciones de los retornos pasados y de la varianza actual estén negativamente correlacionadas (Harvey y Shephard, 1996).

A pesar de que los modelos de volatilidad estocástica son flexibles para modelar los hechos empíricos y sus propiedades estadísticas son más fáciles de obtener que las de los modelos de clase ARCH, su estimación es más complicada, dado que no es posible obtener la función de verosimilitud correspondiente. Esto se debe a que el modelo está caracterizado por una variable latente, la varianza, que no puede ser reconstruida en función de datos observables, como en el caso de los modelos tipo ARCH. Los primeros métodos analíticos fueron cuasi-máxima verosimilitud (QML),

que consiste en aplicar el filtro de Kalman a una transformación del modelo, y el Método Generalizado de Momentos (GMM). Sin embargo, estos métodos tienen problemas de eficiencia en muestras pequeñas, tal como fue demostrado por Jacquier, Polson y Rossi (1994) y Andersen y Sørensen (1996) mediante estudios de simulación para el caso del modelo SV sin efecto leverage.

Con el objetivo de superar los problemas de los métodos analíticos, han surgido métodos que se basan o recurren a simulaciones. Uno de ellos es el Método de Momentos Eficiente (EMM), propuesto por Bansal *et al.* (1995) y aplicado a la estimación de modelos de volatilidad estocástica por Gallant y Tauchen (1996). El EMM es un método que recurre a un modelo auxiliar estimable por QML, por lo general de la familia ARCH. La idea es optimizar un criterio de momentos en el que los momentos son los valores esperados de los scores del modelo auxiliar bajo el modelo estructural de volatilidad estocástica, evaluados en los estimadores QML de los parámetros auxiliares. Debido a que en general no es posible obtener expresiones analíticas para estos valores esperados bajo el modelo de volatilidad estocástica, los momentos del EMM se deben estimar mediante simulaciones. Para el modelo SV sin efecto leverage, Andersen, Chung y Sørensen (1999) y Chumacero (1997) encontraron que el EMM tiene mejor comportamiento en muestras pequeñas que GMM y es una alternativa plausible a otros métodos computacionalmente más demandantes.

Andersen, Chung y Sørensen (1999) estudiaron la eficiencia del EMM para estimar el modelo SV sin efecto leverage, encontrando en general buenos resultados. Un primer intento por establecer la capacidad del EMM para estimar modelos SV con efecto leverage se debe a Giraldo y Lopera (2006), quienes encontraron problemas de convergencia de los estimadores, pero plantean que estos problemas no se deben al método como tal, sino que son causados por el modelo auxiliar utilizado o por el método de optimización. En esta tesis, se intentan superar esos problemas. El objetivo principal es determinar si el EMM es un método eficiente para estimar el modelo SV con efecto leverage y bajo cuáles modelos auxiliares el método es más preciso. Para esto se estudia la eficiencia del método vía simulaciones, recurriendo a diferentes modelos auxiliares. Además de los modelos auxiliares tradicionalmente utilizados, se propone el modelo APARCH de Ding, Granger y Engle (1993), debido a que admite leverage y es flexible, parsimonioso y estable.

La eficiencia del EMM es mayor cuando el modelo auxiliar es una buena aproximación del verdadero proceso generador de datos (Gallant y Long, 1997). Además,

Andersen, Chung y Sørensen (1999) advierten que la inferencia en el EMM es sensible a la elección del modelo auxiliar cuando se tienen muestras pequeñas (menores a 1.000). De aquí que es crucial acertar en la elección de un modelo auxiliar adecuado. No obstante, en la literatura esta elección regularmente se hace, en cierto sentido, de forma arbitraria. De hecho, en este paso es común que se cometan contradicciones como recurrir a modelos auxiliares ARCH y GARCH estacionarios, que no tienen efecto leverage ni memoria larga, para estimar modelos SV con efecto leverage o memoria larga. Al hacer esto, o bien se denota que no se cree que el modelo SV que se está estimando sea acertado o, peor aún, se está abusando del método de estimación. Con el objetivo de dar algunas pistas con respecto a la elección de un modelo auxiliar adecuado, en la sección 1.2 se hace un estudio comparativo de diferentes alternativas, entre ellas el APARCH. Se analiza la precisión de estos modelos para estimar la varianza condicional de procesos SV simulados. Esta comparación permitirá determinar *a priori* cuáles serían los mejores modelos auxiliares para utilizar en la estimación por EMM, de manera que se puedan minimizar los problemas de eficiencia asociados al modelo auxiliar.

En el capítulo 2 se recopilan las principales propiedades estadísticas del modelo SV, se muestra la forma cómo acomoda hechos empíricos de las series de retornos y se presentan algunas extensiones, que tratan de reproducir otras características de los retornos, como colas más pesadas en la distribución de los retornos. El método de momentos eficiente, su fundamento teórico y sus propiedades asintóticas se presentan en el capítulo 3, donde se destaca la importancia de la elección del modelo auxiliar.

El principal aporte de la tesis es el capítulo 4. En este capítulo se hace el estudio de simulaciones del método EMM. Se simulan datos de modelos SV con distribuciones condicionales de los retornos Gaussianas y de colas pesadas, para diferentes tamaños de muestra, con los cuales se hacen las estimaciones por EMM recurriendo a diferentes modelos auxiliares. Se determina para cuáles modelos auxiliares y tamaños de muestra el EMM es más eficiente y se concluye si el EMM es o no un buen método para estimar esta clase de modelos de volatilidad estocástica con efecto leverage.

Finalmente, en el capítulo 5 se presentan las conclusiones y se subrayan los principales aportes de esta investigación.

1.2. Revisión del Problema de Investigación

Sea $y_t \equiv \log P_t/P_{t-1} - \mu$ los retornos diarios de una acción, descontando su valor medio, μ , que se asume constante. El modelo log-normal de volatilidad estocástica (SV) está dado por las ecuaciones:

$$y_t = \sigma_t \varepsilon_t, \quad (1.1)$$

$$h_t = \alpha(1 - \phi) + \phi h_{t-1} + \sigma_\eta \eta_t, \quad (1.2)$$

donde $\sigma_\eta > 0$, $0 < \phi < 1$ y $h_t \equiv \log \sigma_t^2$ es el logaritmo de la varianza condicional de los retornos. Además, se asume que

$$\begin{pmatrix} \varepsilon_{t-1} \\ \eta_t \end{pmatrix} \sim \text{iid} N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right].$$

Note que $\{h_t\}$ sigue un proceso AR(1) con innovaciones Gaussianas. La restricción $0 < \phi < 1$ garantiza que el proceso es estacionario, con media α y varianza finita $\sigma_h^2 = \sigma_\eta^2 / (1 - \phi^2)$. Por propiedades de la distribución log-normal, el proceso de varianza condicional $\{\sigma_t^2\}$ es estacionario con media

$$\mathbb{E}(\sigma_t^2) = \exp \left\{ \alpha + \frac{\sigma_h^2}{2} \right\} \quad (1.3)$$

y varianza

$$\text{Var}(\sigma_t^2) = \exp\{2\alpha + \sigma_h^2\}(\exp\{\sigma_h^2\} - 1). \quad (1.4)$$

Harvey y Shephard (1996) observan que si $\rho = \text{Corr}(\varepsilon_{t-1}, \eta_t)$ es negativo, se obtiene el efecto leverage descrito en la sección 1.1. Yu (2005) muestra que cuando $\rho < 0$, existe una correlación negativa entre $\mathbb{E}(h_{t+1}|y_t)$ y y_t . Esta esperanza condicional permite una interpretación del leverage más clara, ya que muestra que el valor esperado de la volatilidad futura aumenta con una reducción de los retornos en el periodo actual.

Una forma estándar de modelar el efecto leverage es haciendo

$$\eta_t = \rho \varepsilon_{t-1} + \sqrt{1 - \rho^2} \varepsilon_t, \quad (1.5)$$

con $\varepsilon_t \sim \text{iid} N(0, 1)$ independiente de ε_t para todo t . Así, se cumple que $\text{Corr}(\varepsilon_{t-1}, \eta_t) = \rho$ y la ecuación para la log-varianza se transforma en

$$h_t = \alpha(1 - \phi) + \phi h_{t-1} + \sigma_\eta \rho \varepsilon_{t-1} + \sigma_\eta \sqrt{1 - \rho^2} \varepsilon_t. \quad (1.6)$$

Las principales propiedades de este modelo y sus extensiones se describen en el capítulo 2.

Los modelos SV son una clase de modelos parsimoniosos, flexibles para modelar hechos empíricos y con interesantes propiedades estadísticas. Su principal problema consiste en que la estimación del vector de parámetros $\lambda = (\alpha, \phi, \sigma_\eta, \rho)'$ no es sencilla, dado que la función de verosimilitud para una muestra $\{y_t\}_{t=1}^T$,

$$L(\lambda|y) = \int \cdots \int f(\{y_t\}_{t=1}^T | \{h_t\}_{t=1}^T, \lambda) f(\{h_t\}_{t=1}^T | \lambda) dh_1 \cdots dh_T,$$

no está disponible analíticamente, debido a que el proceso $\{h_t\}$ no es observado (Watanabe, 1999). Por esta razón, han aparecido métodos de estimación que no recurren a la verosimilitud. Uno de ellos es el Método de Momentos Eficiente (EMM), propuesto por Gallant y Tauchen (1996) para la estimación de modelos de volatilidad estocástica. Éste es un método indirecto que utiliza los scores de un modelo auxiliar como condiciones de momentos en un criterio GMM. El método se describe con detalle en el capítulo 3. Por ahora, basta con mencionar que en la medida en que el modelo auxiliar se ajusta mejor a los datos, el estimador EMM es más eficiente (Gallant y Tauchen (1996), Gallant y Long (1997)). Adicionalmente, se ha encontrado que la inferencia es sensible a la elección del modelo auxiliar en muestras pequeñas (menores a 1.000), aunque para muestras grandes (mayores a 4.000) el método es robusto a la elección del modelo auxiliar (Andersen, Chung y Sørensen, 1999). Dado que en la práctica no es común disponer de tamaños de muestra tan grandes, resulta crucial la elección de un modelo auxiliar adecuado en la estimación por EMM. Los modelos auxiliares más empleados en la literatura son de la familia ARCH, debido a que son fácilmente estimables por QML, y éste es precisamente el primer paso del método (ver capítulo 3).

Andersen, Chung y Sørensen (1999) estudiaron la precisión del EMM para estimar el modelo SV sin efecto leverage a través de simulaciones y encontraron resultados aceptables, siendo un método similar a otros computacionalmente más demandantes y de implementación más compleja. Pero como ha sido reconocido por Gallant, Hsieh y Tauchen (1997) y Jacquier, Polson y Rossi (2004), un modelo SV plausible para modelar los datos reales debe contener efecto leverage. Por esta razón es importante establecer la precisión del EMM para la estimación de modelos SV que incorporan este efecto. El objetivo principal de este proyecto es determinar si el EMM produce buenos

estimadores para este tipo de modelos y establecer bajo cuáles modelos auxiliares el método es más eficiente.

1.2.1. Comparación de modelos auxiliares

Para la estimación por EMM, usualmente se recurre a modelos auxiliares ARCH, GARCH, EGARCH y a densidades semiparamétricas (SNP), que en su forma más común contienen una componente paramétrica de la familia ARCH para modelar la varianza condicional, y una estructura no paramétrica para distribución de los retornos, la cual recoge otras características de los datos, como colas pesadas y asimetría, que no pueden ser capturadas por una distribución condicional normal (ver, por ejemplo, Andersen, Chung y Sørensen (1999), Gallant, Hsieh y Tauchen (1997) y Gallant y Tauchen (2002)). En esta sección se compara la precisión con la que estos modelos aproximan la varianza generada por los procesos SV. Aquí solamente nos interesa el modelamiento de la varianza condicional, ya que características no Gaussianas de la distribución de los retornos que pueden estar presentes en los datos, pueden ser capturadas agregando términos a la parte no paramétrica de una densidad SNP. La idea es proporcionar pistas acerca de cuál modelo auxiliar podría producir los estimadores EMM más eficientes.

En este punto es importante mencionar que para la estimación por EMM con datos reales, una de las características que debe poder reproducir un buen modelo auxiliar es el efecto leverage. Engle y Ng (1993) comparan varios modelos de la familia ARCH que reproducen el efecto leverage. Denominando como “curva de impacto de noticias” a la relación entre los retornos del día anterior y la varianza actual, encuentran que los modelos EGARCH de Nelson (1991) y GJR-GARCH de Glosten, Jagannathan y Runkle (1993) (en adelante, GJR) son los que mejor se ajustan a las características empíricas de esta curva. Los otros modelos alternativos no ofrecen buenos ajustes. Sin embargo, Engle y Ng (1993) no tienen en cuenta el modelo APARCH de Ding, Granger y Engle (1993), del cual el GJR es un caso particular. Por esta razón, aquí se estudian los modelos EGARCH, GJR y APARCH como posibles modelos auxiliares. El modelo básico GARCH se utiliza como referente para comparar la precisión de los modelos auxiliares con y sin leverage para aproximar la varianza generada por un proceso SV con efecto leverage.

El modelo EGARCH(1,1) está dado por las ecuaciones:

$$\begin{aligned} y_t &= \sigma_t Z_t, & Z_t &\sim \text{iid}(0, 1), \\ \log(\sigma_t^2) &= \omega + \beta \log(\sigma_{t-1}^2) + \theta Z_{t-1} + \gamma(|Z_{t-1}| - \mathbb{E}|Z_{t-1}|), \end{aligned} \quad (1.7)$$

con $\omega, \theta, \gamma \in \mathbb{R}$ y $|\beta| < 1$. Este modelo reproduce el efecto leverage a través de la función asimétrica $g(Z_{t-1}) = \theta Z_{t-1} + \gamma(|Z_{t-1}| - \mathbb{E}|Z_{t-1}|)$. Al igual que en el modelo SV, en el EGARCH(1,1) el logaritmo de la varianza condicional también es un proceso causal AR(1). La diferencia con el modelo SV es la innovación del proceso, ya que mientras en el SV ésta es aleatoria, en el EGARCH(1,1) es una función asimétrica del rezago de la innovación del proceso de retornos. Para el EGARCH(1,1) con $\mathbb{E}(Z_t^2) < \infty$ y $\text{Var}[g(Z_{t-1})] < \infty$, la condición $|\beta| < 1$ es suficiente para la existencia de una única solución estacionaria (He, Teräsvirta y Malmsten, 2002).

El EGARCH tiene algunos problemas. En primer lugar, no existen expresiones sencillas para los momentos del modelo y para las autocorrelaciones de potencias de los retornos absolutos, por lo que en general se deben aproximar numéricamente (ver He, Teräsvirta y Malmsten (2002)). Por otro lado, Straumann y Mikosch (2006) demuestran bajo ciertas condiciones que el estimador QML del EGARCH(1,1) es fuertemente consistente, pero la distribución asintótica del estimador permanece sin ser establecida. Adicionalmente, como veremos más adelante, la recursión para la reconstrucción de la varianza puede ser inestable en algunos casos.

La segunda clase de modelos que se va a considerar es el modelo GARCH asimétrico de potencias o APARCH(1,1), el cual se escribe como

$$\begin{aligned} y_t &= \sigma_t Z_t, & Z_t &\sim \text{iid}(0, 1), \\ \sigma_t^\delta &= \alpha_0 + \alpha(|y_{t-1}| - \gamma y_{t-1})^\delta + \beta \sigma_{t-1}^\delta, \end{aligned} \quad (1.8)$$

con $\alpha_0 > 0$, $\delta \geq 0$, $\beta \geq 0$, $\alpha \geq 0$ y $|\gamma| < 1$. Este modelo también reproduce el efecto leverage, el cual surge cuando $0 < \gamma < 1$. Al permitir que el exponente δ sea un parámetro libre, este modelo impone una transformación de potencias de Box-Cox a la desviación estándar condicional. Esto permite que durante la estimación se obtenga el valor de δ que mejor se ajusta a los datos, haciendo a éste un modelo flexible. Al proponer este modelo, Ding, Granger y Engle (1993) sostienen que puede acomodar memoria larga de potencias de valores absolutos de los retornos, el cual es un hecho empírico frecuentemente observado en las series financieras.

El modelo APARCH incluye como casos especiales otras especificaciones que han sido propuestas en la literatura. Haciendo $\delta = 2$, se obtiene el modelo GJR. Si $\delta = 2$ y $\gamma = 0$, el modelo APARCH se reduce al GARCH(1,1) de Bollerslev (1986), que es un modelo que no incluye el efecto leverage. En Ding, Granger y Engle (1993) se muestran otros modelos como casos particulares. A diferencia del EGARCH, Straumann y Mikosch (2006) demuestran que el estimador QML del GJR es asintóticamente normal.

De acuerdo con Ling y McAleer (2002), la condición suficiente y necesaria para la existencia de una única solución estacionaria de orden δ en el modelo APARCH(1,1) es

$$\alpha \mathbb{E}[(|Z_t| - \gamma Z_t)^\delta] + \beta < 1, \quad (1.9)$$

en cuyo caso y_t es estacionario estricto y ergódico, con $\mathbb{E}(|y_t|^\delta) < \infty$. Note que esta condición depende de la distribución de la innovación $\{Z_t\}$. Ling y McAleer (2002), en el teorema 3.2 muestran las condiciones para la existencia de los momentos de orden $m\delta$ con $m > 0$ del modelo APARCH.

Con el objetivo de comparar la precisión con la que los modelos auxiliares aproximan los modelos SV, se simulan series de datos de modelos SV de un conjunto de parámetros, luego se estiman los modelos auxiliares por QML con estos datos y se reconstruyen las varianzas condicionales. Las especificaciones de los parámetros de los modelos SV utilizados están dadas en términos del cuadrado del coeficiente de variación, el cual para el modelo SV log-normal está dado por

$$CV2 = \frac{\text{Var}(\sigma_t^2)}{[\mathbb{E}(\sigma_t^2)]^2} = \exp\{\sigma_h^2\} - 1, \quad (1.10)$$

donde $\sigma_h^2 = \sigma_\eta^2 / (1 - \phi^2)$. Note que este coeficiente es alto cuando σ_η^2 es alto, es decir, cuando la componente aleatoria de la log-varianza es dominante. Este coeficiente, por lo tanto, se puede interpretar como una medida del grado de persistencia de la varianza condicional en el modelo SV. Jacquier, Polson y Rossi (1994) argumentan que el CV2 es más interpretable que σ_η^2 por sí solo, ya que también involucra el parámetro autorregresivo ϕ . Dado un valor de ϕ , entre menor sea el CV2, mayor es el grado de persistencia de la varianza condicional.

Para el experimento se consideraron 3 grupos de parámetros para el modelo SV, en función del grado de persistencia, los cuales se muestran en la tabla 1.1. Las especificaciones consideradas están basadas en Jacquier, Polson y Rossi (1994), con ligeras

modificaciones. El primer grupo de parámetros está calibrado para un coeficiente de variación al cuadrado de 10. Este caso se denomina de varianza condicional “aleatoria”, dado que el parámetro σ_η es alto y dominante, de modo que los cambios en la varianza condicional son altamente aleatorios. El segundo caso es un conjunto de parámetros tal que se tiene un coeficiente de variación al cuadrado de 2. Esta especificación se denomina de varianza condicional “persistente”, dado que el parámetro autorregresivo ϕ es 0.95 y σ_η tiene un valor intermedio. La varianza “altamente persistente” es el tercer caso, con un CV2 de 0.1, donde σ_η es pequeño y ϕ es 0.98, de manera que la varianza condicional está altamente autocorrelacionada. Todas las especificaciones están calibradas para un valor esperado de la varianza condicional de 0.0007.

Varianza	Parámetros				$\mathbb{E}(\sigma_t^2)$	$\text{Var}(\sigma_t^2)$	CV2
	α	ϕ	σ_η	ρ			
1. Aleatoria	-8.450	0.80	0.93000	-0.7	0.0007	5.08×10^{-6}	10
2. Persistente	7.810	0.95	0.32730	-0.7	0.0007	9.88×10^{-7}	2.0
3. Altamente persistente	-7.312	0.98	0.06145	-0.7	0.0007	4.90×10^{-8}	0.1

Tabla 1.1: Parámetros de los modelos de volatilidad estocástica simulados. Todas las especificaciones están calibradas para una varianza condicional media de 0.0007.

En primer lugar, se generaron 10.000 datos de los modelos SV y con estos se estimaron los modelos auxiliares GARCH, EGARCH, GJR y APARCH por QML. Para comparar la precisión de las estimaciones de la varianza condicional, se calcularon los promedios acumulados de las varianzas estimadas por cada modelo, con el fin de observar a qué valores convergen las varianzas promedio. En la figura 1.1 se muestran los resultados para la especificación de modelo SV con varianza condicional “aleatoria”. Se aprecia que el promedio de la varianza condicional estimada con GJR tiende a converger a un valor que duplica el del modelo SV. El modelo GARCH también genera una varianza condicional promedio estimada muy alta en comparación con la verdadera. En este caso, los modelos que más se aproximan son el EGARCH, cuyas varianzas promedio son idénticas a las del SV, y el APARCH, que tiene varianzas condicionales promedio ligeramente superiores.

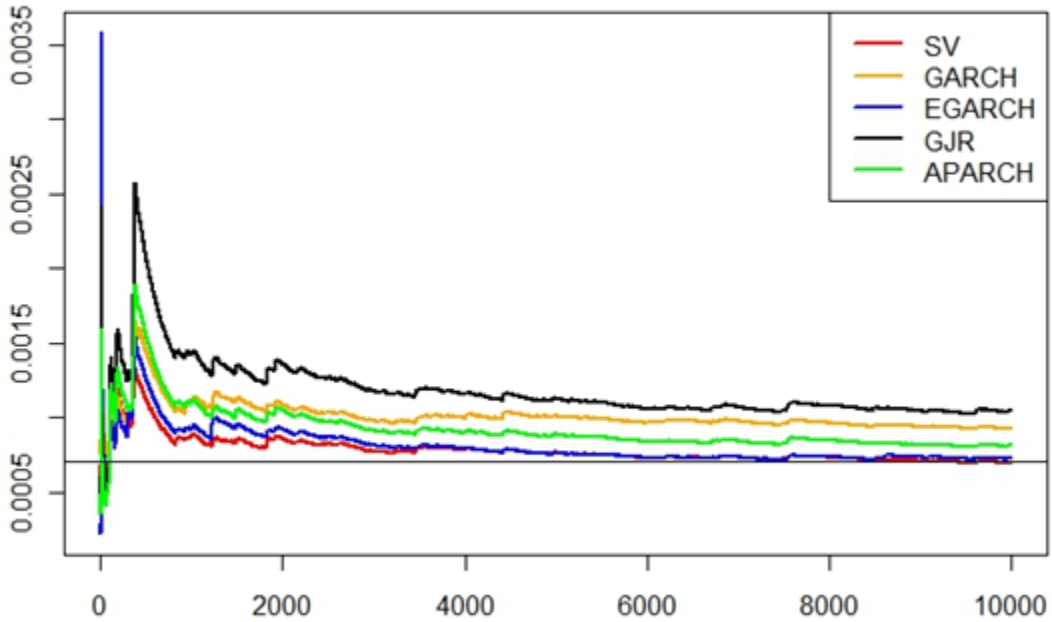
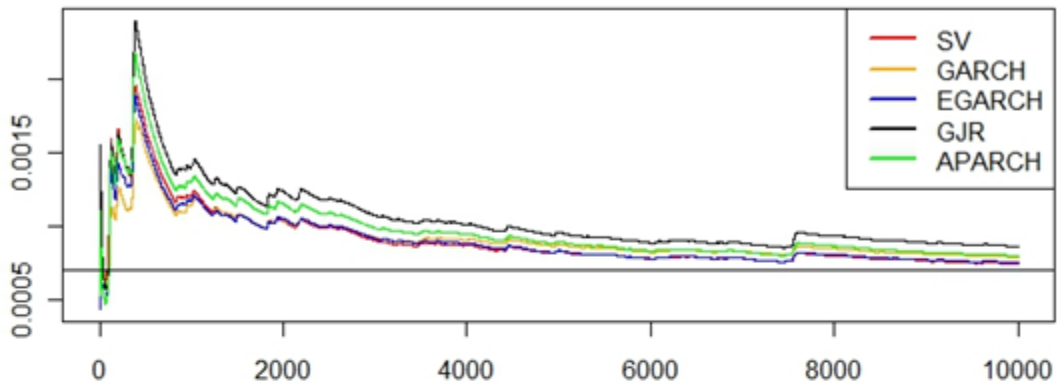
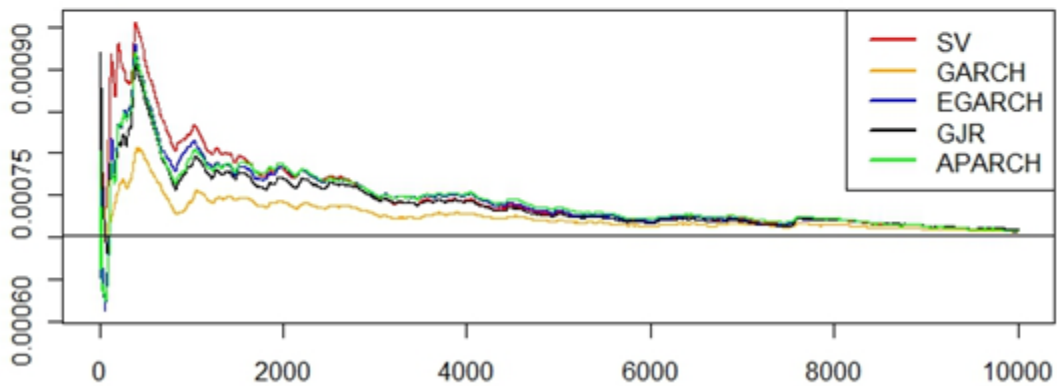


Figura 1.1: Promedio acumulado de la varianza condicional estimada para 10.000 datos simulados del modelo de volatilidad estocástica con varianza condicional “aleatoria” (ver tabla 1.1). La línea horizontal corresponde al valor esperado de la varianza condicional del modelo SV, que es 0.0007.

En la figura 1.2 se presentan los dos casos con datos simulados de procesos SV con varianza condicional “persistente” y “altamente persistente”. En comparación con el caso anterior, es claro que a medida que la varianza condicional es más persistente, las varianzas promedio estimadas tienden a converger a valores más cercanos a los del modelo SV. No obstante, aún con varianza persistente, la varianza promedio estimada del modelo GJR sigue siendo muy alta en comparación con la verdadera. Cuando la varianza es altamente persistente, todos los modelos auxiliares ofrecen resultados similares. Hasta aquí hemos visto que, en promedio, los modelos auxiliares GJR y GARCH tienden a sobrestimar considerablemente la varianza condicional de los modelos SV cuando la varianza es aleatoria y que las estimaciones mejoran en la medida que la varianza condicional es más persistente.



(a) Varianza persistente



(b) Varianza altamente persistente

Figura 1.2: Promedio acumulado de la varianza condicional estimada para 10.000 datos simulados de modelos de volatilidad estocástica con: (a) varianza condicional “persistente”, y (b) varianza condicional “altamente persistente” (ver tabla 1.1). La línea horizontal corresponde al valor esperado de la varianza condicional del modelo SV, que es 0.0007.

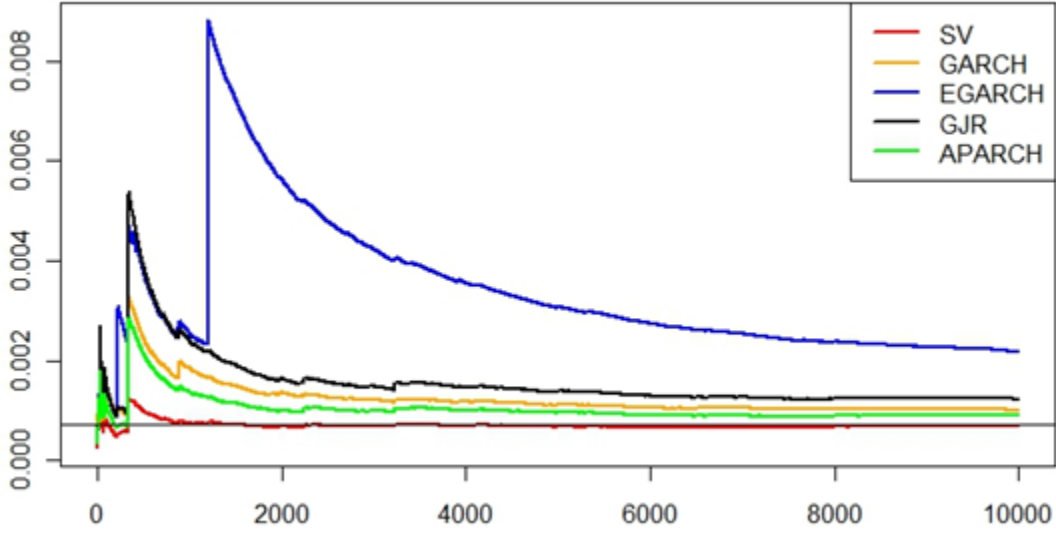


Figura 1.3: Inestabilidad del EGARCH. Una simulación y estimación de la varianza para 10.000 datos de un modelo de volatilidad estocástica con varianza condicional “aleatoria” (ver tabla 1.1). La línea horizontal corresponde al valor esperado de la varianza condicional del modelo SV, que es 0.0007.

A pesar de que en general se obtienen buenos resultados con el modelo EGARCH, se debe tener en cuenta que este proceso ocasionalmente se vuelve inestable. El EGARCH(1,1) obedece la ecuación recursiva

$$\log(\sigma_t^2) = \omega - \gamma \mathbb{E}|Z_{t-1}| + \beta \log(\sigma_{t-1}^2) + e^{-(\log \sigma_{t-1}^2)/2} (\theta y_{t-1} + \gamma |y_{t-1}|), \quad (1.11)$$

la cual, en determinados casos, puede salirse de control y generar valores extremadamente altos de la varianza condicional. Debido a la parte exponencial, es posible pasar de un valor negativo muy grande a un valor positivo muy grande de la log-varianza. Esta inestabilidad se presenta de forma imprevisible, dado que puede ocurrir entre simulaciones de un mismo proceso. En la figura 1.3 se muestra un caso en el que el EGARCH es inestable, el cual aparece en una simulación del proceso SV con varianza condicional aleatoria. En ella se observa que ocurre una estimación de la varianza condicional extremadamente grande, que causa que el promedio de las varianzas estimadas se mantenga muy alto y no tienda a valores cercanos a los del modelo SV. Por

el contrario, las estimaciones GARCH, GJR y APARCH son consistentes y estables en cada simulación.

Este resultado es consistente con los hallazgos de Engle y Ng (1993), quienes también encuentran que grandes valores negativos de los retornos hacen que el modelo EGARCH pronostique varianzas condicionales que son hasta 3.000 veces mayores a las observadas. Esto genera que la desviación estándar de la varianza condicional sea considerablemente mayor a la implicada por modelos GARCH y GJR y también es mayor a la de los retornos al cuadrado. Por ende, el EGARCH puede generar una distribución de la varianza condicional con colas demasiado extremas. Por otro lado, Engle y Ng (1993) también hallan que en periodos de estabilidad, bajo el modelo EGARCH se tiende a rechazar la hipótesis de efecto leverage, aún cuando éste se encuentra presente. Por el contrario, el modelo GJR tiende a no rechazar esta hipótesis. Engle y Ng (1993) argumentan que esto se debe a que la forma cuadrática del GJR domina a la forma exponencial del EGARCH con valores pequeños de los retornos. Esto implica que un retorno negativo, pero pequeño, no afecta de forma significativa a la varianza condicional del siguiente periodo en el EGARCH, lo que hace indetectable el efecto leverage.

El problema de una varianza explosiva es que afecta el algoritmo de optimización para la estimación por EMM, por lo que se requiere tener precaución cuando se utiliza el EGARCH como modelo auxiliar. Una posible explicación de la inestabilidad del EGARCH podría encontrarse en lo que Straumann y Mikosch (2006) denominan propiedad de invertibilidad de los modelos de heteroscedasticidad condicional, que no es más que la convergencia en probabilidad del algoritmo estimador de la varianza a medida que el tamaño de la muestra va a infinito. Para el EGARCH(1,1), esta propiedad depende de los parámetros del modelo (ver Straumann y Mikosch (2006), proposición 3.9), mientras que para el GJR(1,1), la estacionariedad garantiza que el proceso es invertible. Por esta razón, ciertas combinaciones de parámetros estimados en el EGARCH(1,1) podrían conducir a que la reconstrucción de la varianza se salga de control.

Con el propósito de cuantificar la forma promedio en que los modelos auxiliares estiman la verdadera volatilidad condicional, se generaron 1.000 simulaciones para tamaños de muestra de 1.500 de cada una de las especificaciones de modelos SV de la tabla 1.1. Para cada una de estas series se estimaron las volatilidades con los modelos auxiliares ($\hat{\sigma}_t$) y se obtuvieron las frecuencias de las desviaciones porcentuales de

las estimaciones de la volatilidad, $(\hat{\sigma}_t - \sigma_t)/\sigma_t \times 100\%$, donde σ_t son las verdaderas volatilidades generadas por los modelos SV. Estas desviaciones se clasificaron en 5 intervalos: i) estimaciones por debajo del 50% del valor real de la volatilidad, ii) estimaciones entre 50% y 20% menores al valor real, iii) estimaciones entre 20% menores y 20% mayores al valor real, iv) estimaciones entre 20% y 50% mayores a la verdadera volatilidad y v) estimaciones de volatilidades que exceden en más del 50% a los valores reales. Luego se hicieron los promedios de estas frecuencias. Evidentemente, un buen modelo auxiliar es aquél que genera una proporción alta de estimaciones de la volatilidad condicional con magnitudes que no exceden ni son inferiores en un 20% a los verdaderos valores. En la tabla 1.2 se presentan los resultados.

Como se puede observar en el panel A de la tabla 1.2, las volatilidades condicionales son altamente sobrestimadas por los modelos auxiliares cuando la varianza es aleatoria, generando valores mayores al 50% de los verdaderos más del 20% de las veces, y con el modelo GARCH teniendo el peor desempeño. Los modelos que mejor se comportan son el EGARCH y el APARCH que generan el 43.1% y el 41.7% de las estimaciones en el intervalo central, respectivamente. En el panel B, con varianza condicional persistente, las estimaciones mejoran para todos los modelos auxiliares. En este caso, los modelos EGARCH, APARCH y GJR generan estimaciones en el intervalo central más del 60% de las veces. No obstante, aún estos modelos sobrestiman las verdaderas volatilidades con magnitudes entre 20% y 50% mayores a las verdaderas, con frecuencias cercanas al 20%. En el extremo de varianza condicional altamente persistente (panel C), las estimaciones con los modelos EGARCH, APARCH y GJR tienen una precisión superior al 95%. El modelo GARCH, por el contrario, sigue sobrestimando la verdadera volatilidad el 11.9% de las veces.

Modelo auxiliar	Frecuencias relativas(%)				
	i	ii	iii	iv	v
	<-50 %	(-50 %,-20 %)	(-20 %,20 %)	(20 %,50 %)	>50 %
A. Varianza condicional “aleatoria”					
GARCH	1.8	17.9	31.0	13.1	36.3
EGARCH	0.9	14.9	43.1	20.0	21.2
GJR	1.3	15.5	38.8	19.9	24.5
APARCH	1.2	15.2	41.7	20.7	21.3
B. Varianza condicional “persistente”					
GARCH	0.2	16.1	46.1	19.0	18.6
EGARCH	0.1	9.3	68.2	18.5	4.0
GJR	0.1	11.6	60.5	20.6	7.2
APARCH	0.1	10.1	66.0	19.2	4.7
C. Varianza condicional “altamente persistente”					
GARCH	0.0	3.0	85.2	11.4	0.5
EGARCH	0.0	0.6	97.4	2.0	0.0
GJR	0.0	0.6	95.2	4.2	0.0
APARCH	0.0	0.4	97.4	2.2	0.0
D. Modelo SV sin efecto leverage					
GARCH	1.0	14.8	41.4	23.0	19.7
EGARCH	1.0	14.9	41.6	23.0	19.5
GJR	1.1	14.9	41.4	23.0	19.7
APARCH	1.0	14.9	41.6	23.0	19.5

Tabla 1.2: Distribución de frecuencias de las desviaciones porcentuales de las volatilidades estimadas mediante modelos auxiliares con respecto a las verdaderas volatilidades condicionales generadas por los modelos de volatilidad estocástica. Se promedian 1.000 repeticiones para tamaños de muestra de 1.500. Para los parámetros utilizados en las simulaciones, ver tabla 1.1. El panel D, modelo sin efecto leverage, corresponde a los mismos parámetros del modelo SV con varianza condicional “persistente”, excepto que el parámetro de leverage, ρ , se hace igual a cero.

Por otro lado, cabe resaltar que las estimaciones de δ de los modelos APARCH son menores al valor de 2 asumido por el modelo GJR (no reportado), lo cual puede

explicar por qué la recursión cuadrática del GJR tiende a sacar valores altos de la varianza condicional. Este resultado es consistente con el hallazgo de Engle y Ng (1993), quienes también encuentran que el GJR puede generar valores muy altos de la varianza condicional, cuando hay datos extremos de los retornos. El modelo APARCH es, sin duda, superior al modelo GJR.

Para la especificación sin leverage (panel D), que tiene los mismos parámetros del modelo persistente pero con $\rho = 0$, todos los modelos son tan efectivos como el GARCH. Este resultado es acorde con los hallazgos de Andersen, Chung y Sørensen (1999), quienes para la estimación por EMM de modelos SV sin efecto leverage, encuentran mejores resultados cuando se emplean modelos auxiliares GARCH y EGARCH simétricos (con parámetro de asimetría igual a cero), que cuando se utiliza el EGARCH asimétrico. Un resultado interesante en este caso es que se observa que las estimaciones tienden a sobrestimar la volatilidad más que en el caso con efecto leverage (compare paneles B y D). Una posible explicación a esto es que el efecto leverage introduce cierto grado de predicción de la varianza que mejora las estimaciones.

1.2.2. Síntesis

El principal resultado de este estudio comparativo es la observación de que cuando la varianza condicional es altamente aleatoria, los modelos auxiliares de clase ARCH tienden a sobrestimar las verdaderas volatilidades generadas por los modelos SV, pero también se vio que las estimaciones mejoran notablemente cuando la varianza condicional es más persistente. Los modelos EGARCH y APARCH ofrecen las mejores estimaciones, mientras que el modelo GARCH es, de lejos, el que peor ajuste ofrece en todos los casos. Este modelo tiende a sobrestimar considerablemente la verdadera varianza condicional, resultado que era de esperar, puesto que este modelo no reproduce el efecto leverage. El EGARCH es un buen modelo auxiliar, pero hay que tener precauciones, dado que la ecuación recursiva puede salirse de control en algunos casos. Por su parte, el APARCH se muestra como un modelo competitivo y estable.

Estos resultados confirman que es importante utilizar modelos auxiliares EGARCH o APARCH en la estimación por EMM cuando se tienen datos con efecto leverage. Los autores que no tienen esto en cuenta, pueden no encontrar buenos resultados en la estimación de modelos SV por EMM, pero la falencia puede no estar en el modelo SV necesariamente, sino en la elección del modelo auxiliar.

Capítulo 2

El Modelo de Volatilidad Estocástica

2.1. Introducción

Recordemos de la sección 1.2 que el modelo de volatilidad estocástica log-normal está dado por las ecuaciones:

$$y_t = \sigma_t \varepsilon_t, \quad (2.1)$$

$$h_t = \alpha(1 - \phi) + \phi h_{t-1} + \sigma_\eta \eta_t, \quad (2.2)$$

donde

$$\begin{pmatrix} \varepsilon_{t-1} \\ \eta_t \end{pmatrix} \sim \text{iid} N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right],$$

con los parámetros $\sigma_\eta > 0$, $0 < \phi < 1$ y $h_t \equiv \log(\sigma_t^2)$. Además, recordemos que el efecto leverage aparece cuando $\rho < 0$. Como vimos en la sección 1.1, esta clase de modelos surgió para tratar de explicar los hechos empíricos de las series de retornos financieros. Según Shephard y Andersen (2009), Jungbacker y Koopman (2009) y Yu (2005), el modelo de volatilidad estocástica log-normal puede interpretarse como la versión discreta de un proceso en tiempo continuo en el cual el logaritmo de los precios

del activo (ignorando términos de *drift* y primas de riesgo) y su correspondiente log-varianza siguen un sistema de ecuaciones diferenciales estocásticas de la forma:

$$d \log P_t = \sigma_t dW_t, \quad (2.3)$$

$$d \log \sigma_t^2 = \varphi(\alpha - \log \sigma_t^2) dt + \omega dB_t, \quad (2.4)$$

donde $\varphi > 0$, $\omega > 0$, y W y B son movimientos Brownianos estándar, con correlación ρ . De manera análoga al modelo discreto, cuando $\rho < 0$ surge el efecto leverage. Este modelo fue utilizado por Wiggins (1987) y Melino y Turnbull (1990) para la valoración de opciones financieras. Si bien en la actualidad, los modelos en tiempo continuo han ido ganando terreno dentro de la literatura financiera, los modelos SV en tiempo discreto son preferibles en situaciones en las que se dispone solamente de una observación diaria en lugar de datos de alta frecuencia, esenciales para la estimación precisa de modelos en tiempo continuo.

Los modelos de volatilidad estocástica y la familia de modelos de clase ARCH son las principales alternativas para modelar la varianza condicional en series de tiempo financieras. La diferencia fundamental entre estas clases de modelos es que en los ARCH la volatilidad se genera endógenamente (como resultado de cambios en los retornos pasados, por ejemplo), mientras que en los modelos SV, la ecuación de la varianza condicional contiene un término estocástico que pretende explicar cambios en la volatilidad que en la práctica se asocian principalmente al flujo exógeno y aleatorio de información relevante para la formación de los precios de los activos.

Aparte de esta diferencia conceptual, existen divergencias en cuanto a la flexibilidad de estos modelos para adaptarse a los hechos observados de las series de retornos financieros. Por ejemplo, Jacquier, Polson y Rossi (1994) encontraron evidencia de que la FAC del modelo SV se ajusta mejor que la de un modelo GARCH a la FAC muestral de los retornos al cuadrado de un portafolio de acciones. Por su parte, Carnero, Peña y Ruiz (2004) señalan que el modelo de volatilidad estocástica es más flexible para acomodar varios hechos empíricos simultáneamente, como curtosis alta, colas pesadas, persistencia de la volatilidad y autocorrelaciones de los retornos al cuadrado. Por ejemplo, en el modelo GARCH(1,1) Gaussiano, en el que la varianza condicional está dada por $\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2$, la condición para que el cuarto momento de los retornos sea finito es $(\alpha_1 + \beta_1)^2 + 2\alpha_1^2 < 1$ (Bollerslev, 1986), de manera que cuando la persistencia es alta ($\alpha_1 + \beta_1$ es cercano a 1), este modelo no puede tener curtosis finita. En este mismo sentido, Kim, Shephard y Chib (1998) encuentran

que el modelo SV Gaussiano es mejor que el GARCH Gaussiano y es similar a un modelo t -GARCH cuando se ajusta a datos de tasas de cambio. No obstante, Carnero, Peña y Ruiz (2004) encuentran que aún bajo una distribución de cola pesada, para obtener un modelo GARCH(1,1) persistente y con curtosis alta, pero finita, se debe restringir el parámetro α_1 a cero. En contraste, los modelos SV no imponen esta clase de restricciones sobre los parámetros, como veremos más adelante, de suerte que un modelo SV Gaussiano puede generar series con curtosis alta, sin implicar alta persistencia y sin forzar a ε_t a tener una distribución de cola pesada. Carnero, Peña y Ruiz (2004) muestran que estos resultados también son válidos cuando se comparan los modelos SV con efecto leverage y los modelos de la familia ARCH con efecto leverage (EGARCH y QGARCH, específicamente).

Otro problema de los modelos GARCH es que debido a la forma funcional de la varianza condicional, los valores grandes del proceso tienden a agruparse, lo cual puede no representar fielmente lo que ocurre en la práctica (Davis y Mikosch, 2009a). Adicionalmente, los modelos SV tienen la ventaja de que las expresiones de las FAC de las potencias de los retornos absolutos están disponibles, lo que permite contrastar con la evidencia muestral, mientras que en los modelos GARCH son desconocidas (Davis y Mikosch, 2009b). Por último, cabe señalar que, tal como vimos en la sección 1.2, los modelos de heteroscedasticidad condicional no se ajustan bien a series cuya varianza es altamente aleatoria.

Las propiedades estadísticas de los modelos SV discretos han sido compiladas por Ghysels, Harvey y Renault (1996), Taylor (1994), Shephard y Andersen (2009) y Davis y Mikosch (2009b), entre otros. En la sección 2.2 se recogen las principales propiedades, donde se muestra la flexibilidad del modelo para incluir hechos observados en las series financieras y en la sección 2.3 se presentan algunas importantes extensiones que tratan de reproducir características de los retornos y la varianza que no pueden ser capturadas por el modelo básico.

2.2. Propiedades Estadísticas

Los modelos SV tratan de explicar el fenómeno de los clústeres de la volatilidad, los cuales se capturan al permitir que el proceso de log-varianza condicional esté autocorrelacionado. En el modelo SV básico, las autocorrelaciones son las de un proceso

AR(1), aunque es posible acomodar otra clase de especificaciones, como veremos más adelante.

En este modelo, los retornos son una diferencia martingala, aún con efecto leverage (Harvey y Shephard (1996), Yu (2005)), lo cual es consistente con la hipótesis de mercados eficientes. Sin embargo, como anotan Shephard y Andersen (2009), al no tener en cuenta las primas de riesgo, el modelo básico solo modela la parte riesgosa de los rendimientos, y es esta componente la que no es pronosticable. En situaciones o mercados en los que sea posible introducir una parte pronosticable a los rendimientos, esto se hace fácilmente agregando una componente que contenga la prima de riesgo.

La estructura probabilística del modelo está determinada por las propiedades de los procesos $\{\varepsilon_t\}$ y $\{\sigma_t\}$. Si $\{\varepsilon_t\}$ y $\{\sigma_t\}$ son estrictamente estacionarios, entonces $\{y_t\}$ es estrictamente estacionario (Davis y Mikosch, 2009b). Como señalan Davis y Mikosch (2009b), cualquier proceso $\{h_t\}$ es estacionario si tiene representación lineal causal

$$h_t = \sum_{j=0}^{\infty} \psi_j \eta_{t-j}^*, \quad t \in \mathbb{Z}, \quad (2.5)$$

con $\psi_0 = 1$, $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ y donde η_t^* es una secuencia de variables aleatorias iid de media cero y varianza finita σ_η^2 (ver Brockwell y Davis (1991)). Algunas clases de procesos que cumplen esta propiedad son los ARMA estacionarios, incluyendo el AR(1), y los ARFIMA de memoria larga. Esta propiedad es heredada por funciones medibles de h_t , específicamente σ_t . Es decir, si $\{h_t\}$ es estacionario estricto, entonces $\{\sigma_t\}$ también lo es. Por lo tanto, la condición $0 < \phi < 1$ garantiza que los procesos de volatilidad estocástica y de retornos son estacionarios estrictos en el modelo básico. Por teorema ergódico (Grimmett y Stirzaker (2001), capítulo 9), la estacionariedad implica que $\{y_t\}$ es ergódico, es decir, que su media muestral obedece la ley fuerte de grandes números.

Por independencia contemporánea entre $\{\sigma_t\}$ y $\{\varepsilon_t\}$, los momentos del modelo están dados por el producto de los momentos de estos procesos:

$$\mathbb{E}(y_t^c) = \mathbb{E}(\sigma_t^c) \mathbb{E}(\varepsilon_t^c), \quad c > 0. \quad (2.6)$$

Entonces, la distribución incondicional de los retornos tiene media cero y varianza

$$\text{Var}(y_t) = \mathbb{E}(\sigma_t^2) = \exp \left\{ \alpha + \frac{\sigma_h^2}{2} \right\}, \quad (2.7)$$

donde σ_h^2 es la varianza de h_t , que en el caso AR(1) está dada por $\sigma_\eta^2/(1 - \phi^2)$. Esta varianza es constante. Note que si la distribución de ε_t es simétrica, los momentos impares de y_t son cero. Expresiones de los momentos para el modelo sin efecto leverage se encuentran en Ghysels, Harvey y Renault (1996) y Harvey (1998). Para el modelo con leverage, los momentos son los mismos que para el modelo sin leverage, excepto aquellos que involucran rezagos de los retornos (Harvey y Shephard, 1996).

Debido a que los $\{\varepsilon_t\}$ son incorrelacionados, los retornos son ruido blanco, pero no son independientes, ya que potencias de los retornos absolutos son autocorrelacionados. Para el modelo sin leverage, las autocorrelaciones de las potencias de los retornos absolutos, $|y_t|^c$ para $c > 0$, están dadas por (Harvey, 1998):

$$\rho_c(\tau) = \frac{\exp\left\{\frac{c^2}{4}\sigma_h^2\rho_\tau\right\} - 1}{\kappa_c \exp\left\{\frac{c^2}{4}\sigma_h^2\right\} - 1}, \quad \tau \geq 1, \quad (2.8)$$

donde

$$\kappa_c = \frac{\mathbb{E}(|\varepsilon_t|^{2c})}{[\mathbb{E}(|\varepsilon_t|^c)]^2}.$$

En la anterior expresión, ρ_t denota la FAC de $\{h_t\}$, que en el caso AR(1) está dada por $\rho_t = \phi^t$. Estas autocorrelaciones son válidas para cualquier modelo SV en el que $\{h_t\}$ es un proceso Gaussiano. Taylor (1986) y Harvey (1998) observan que si σ_h^2 es pequeño o ϕ es cercano a uno, entonces

$$\rho_c(\tau) \approx \rho_\tau \frac{\exp\left\{\frac{c^2}{4}\sigma_h^2\right\} - 1}{\kappa_c \exp\left\{\frac{c^2}{4}\sigma_h^2\right\} - 1}, \quad \tau \geq 1. \quad (2.9)$$

Es decir, las autocorrelaciones de $\{|y_t|^c\}$ tienen un comportamiento similar a las de $\{h_t\}$, excepto que están multiplicadas por un factor de proporcionalidad. Bajo la especificación AR(1), Taylor (1986) argumenta que las autocorrelaciones $\rho_c(\tau)$ son similares a las de un proceso ARMA(1,1).

Para los retornos al cuadrado se tiene que $\kappa_c = \kappa_\varepsilon$, donde κ_ε es la curtosis de ε_t (es igual a 3 si se asume que el proceso $\{\varepsilon_t\}$ se distribuye normal), entonces las autocorrelaciones de los retornos al cuadrado están dadas por

$$\rho_2(\tau) = \frac{\exp\{\sigma_h^2\rho_\tau\} - 1}{\kappa_\varepsilon \exp\{\sigma_h^2\} - 1}, \quad \tau \geq 1. \quad (2.10)$$

Dado que $\{\sigma_t\}$ y $\{y_t\}$ son estacionarios ergódicos, las autocorrelaciones muestrales convergen casi seguramente (a.s.) a sus valores poblacionales, siempre que existan y sean finitas (Davis y Mikosch, 2009b).

Para el modelo con efecto leverage, es posible obtener las autocorrelaciones de los retornos al cuadrado cuando $\varepsilon_t \sim N(0, 1)$ y $\{h_t\}$ sigue un proceso AR(1) Gaussiano. En este caso, se tiene que

$$\rho_2(\tau) = \frac{\text{Cov}(y_t^2, y_{t-\tau}^2)}{\text{Var}(y_t^2)} = \frac{\mathbb{E}(\varepsilon_{t-\tau}^2 \sigma_t^2 \sigma_{t-\tau}^2) - [\mathbb{E}(\sigma_t^2)]^2}{3\mathbb{E}(\sigma_t^4) - [\mathbb{E}(\sigma_t^2)]^2}.$$

Las esperanzas $\mathbb{E}(\sigma_t^2)$ y $\mathbb{E}(\sigma_t^4)$ se obtienen haciendo uso de las propiedades de la distribución log-normal, mientras que $\mathbb{E}(\varepsilon_{t-\tau}^2 \sigma_t^2 \sigma_{t-\tau}^2)$ se puede obtener a través de

$$\mathbb{E}(\varepsilon_{t-\tau}^2 \sigma_t^2 \sigma_{t-\tau}^2) = \mathbb{E}[\mathbb{E}(\varepsilon_{t-\tau}^2 \sigma_t^2 \sigma_{t-\tau}^2 | \Omega_{t-\tau})] = \mathbb{E}[\varepsilon_{t-\tau}^2 \sigma_{t-\tau}^2 \mathbb{E}(\sigma_t^2 | \Omega_{t-\tau})],$$

donde $\Omega_{t-\tau}$ es el conjunto de información hasta $(t - \tau)$. Note que $\varepsilon_{t-\tau}^2$ y $\sigma_{t-\tau}^2$ hacen parte del conjunto $\Omega_{t-\tau}$, y por ende pueden sacarse de la esperanza condicional en la anterior expresión. Ahora, para obtener $\mathbb{E}(\sigma_t^2 | \Omega_{t-\tau})$, se expande σ_t^2 hasta τ periodos atrás:

$$\sigma_t^2 = \exp \left\{ \alpha(1 - \phi) \sum_{j=0}^{\tau-1} \phi^j + 2\phi^\tau \log \sigma_{t-\tau} + \sigma_\eta \sum_{j=0}^{\tau-1} \phi^j \eta_{t-j} \right\}.$$

Luego se hace uso de $\eta_t = \rho\varepsilon_{t-1} + \sqrt{1 - \rho^2}\varepsilon_t$, donde ε_t y ε_{t-1} son v.a. normales estándar independientes para obtener $\mathbb{E}(\sigma_t^2 | \Omega_{t-\tau})$. Después de unos cálculos algo tediosos, se llega a la siguiente expresión para la FAC de los retornos al cuadrado en el modelo con efecto leverage:

$$\rho_2(\tau) = \frac{(\rho^2 \sigma_\eta^2 \phi^{2\tau-2} + 1) \exp\{\sigma_h^2 \phi^\tau\} - 1}{3 \exp\{\sigma_h^2\} - 1}, \quad \tau \geq 1. \quad (2.11)$$

Note que esta expresión difiere solo ligeramente de la FAC del modelo sin efecto leverage, ya que $\rho^2 \sigma_\eta^2 \phi^{2\tau-2}$ es un número pequeño. Además, ambas FAC tienen un comportamiento asintótico similar.

Una crítica que podría hacerse al modelo SV básico es el supuesto de que la varianza condicional sigue un proceso log-normal. Sin embargo, este supuesto no solo es conveniente desde el punto de vista algebraico, sino que ha sido defendido desde la

evidencia empírica de la “volatilidad realizada”, un término que denota una estimación de la volatilidad diaria usando datos de alta frecuencia. Suponga que el logaritmo de los precios (en tiempo continuo) es la solución a la ecuación diferencial estocástica (ignorando términos de *drift* y primas de riesgo):

$$d \log P_t = \sigma_t dW_t, \quad (2.12)$$

donde W es un movimiento Browniano estándar. La variación cuadrática (QV) se define como la volatilidad integrada en un periodo de tiempo (Andersen, Bollerslev, Diebold y Labys, 2001):

$$QV(t) = \int_{t-1}^t \sigma_s^2 ds, \quad (2.13)$$

la cual tiene un rol fundamental en los modelos de volatilidad estocástica en tiempo continuo. Una forma de estimar la variación cuadrática diaria es a través de la suma de los cuadrados de las diferencias de los logaritmos de los precios de M observaciones intra-día, muestreadas cada Δ unidades de tiempo, es decir:

$$\{Q\}_t = \sum_{j=1}^M \left[y^* \left(t - 1 + \frac{\Delta j}{M} \right) - y^* \left(t - 1 + \frac{\Delta(j-1)}{M} \right) \right]^2, \quad (2.14)$$

donde $y^*(h) = \log P_h$. Esta variable converge en probabilidad a $QV(t)$ a medida que el ancho del intervalo de muestreo, Δ , va a cero (Barndorff-Nielsen y Shephard (2002) y Andersen, Bollerslev, Diebold y Labys (2001)). $\{Q\}_t$ se conoce como varianza realizada y su raíz cuadrada es la volatilidad realizada. Esta medida permite “observar” con buena precisión el proceso de la volatilidad latente diaria.

Andersen, Bollerslev, Diebold y Ebens (2001) encuentran que el logaritmo de la volatilidad realizada diaria de los retornos de acciones tiene las características de la distribución normal, lo que concede legitimidad al modelo SV log-normal. Estos resultados también son válidos para datos de tasas de cambio, tal como es reportado por Andersen, Bollerslev, Diebold y Labys (2001). Barndorff-Nielsen y Shephard (2002) encuentran que la distribución log-normal o la inversa Gaussiana (NIG) pueden ajustarse bien a las características de la volatilidad realizada, pero además muestran que el modelo log-normal sigue siendo válido aún cuando el número de observaciones intra-día disponibles es pequeño.

En el modelo SV log-normal, la curtosis de los retornos es $\kappa_\varepsilon \exp\{\sigma_h^2\} > \kappa_\varepsilon$. Esta curtosis es mayor a 3 aún asumiendo que $\{\varepsilon_t\}$ es Gaussiano. Por lo tanto, el modelo produce una distribución incondicional estacionaria y leptocúrtica de los retornos, lo cual es consistente con la evidencia empírica. Una forma de acomodar una curtosis más alta es asumiendo una distribución leptocúrtica de la innovación ε_t . Pero nótese que la curtosis también es mayor a medida que la varianza del proceso de la log-varianza condicional (σ_h^2) es más alta. Esto es válido no solo para el modelo log-normal, sino en el caso general (Renault, 2009), y ofrece una forma de probar empíricamente la validez de los modelos SV, contrastando las características de las distribuciones de los retornos con las de la volatilidad realizada diaria de diferentes acciones.

2.3. Extensiones del Modelo Básico

Una de las principales ventajas del modelo SV es su flexibilidad para acomodar hechos empíricos. En algunas series de retornos se ha observado que estos pueden tener autocorrelaciones pequeñas, pero significativas (Taylor (1982, 1986)). Para modelar este hecho, Taylor (1982) asume que el proceso $\{\varepsilon_t\}$ no es una secuencia iid, sino un proceso autocorrelacionado. Pero actualmente es más común asumir que los $\{\varepsilon_t\}$ son iid (Shephard y Andersen, 2009), y las autocorrelaciones se consiguen modelar incluyendo una estructura ARMA(p,q) en el proceso de retornos. Por ejemplo, Gallant, Hsieh y Tauchen (1997) modelan los retornos y la log-varianza como procesos AR(p), con $p > 1$, de la siguiente manera:

$$y_t = \sum_{j=1}^{L_y} c_j y_{t-j} + \sigma_t \varepsilon_t, \quad (2.15)$$

$$h_t - \alpha = \sum_{j=1}^{L_h} \phi_j (h_{t-j} - \alpha) + \sigma_\eta \eta_t. \quad (2.16)$$

Esta especificación permite modelar cambios en la media condicional de y_t , que pueden ocurrir por periodos de altos y bajos retornos, y también permite modelar una estructura de autocorrelaciones de la log-varianza un poco más generalizada que la impuesta por el modelo AR(1). En Harvey y Shephard (1996) se fija $L_y = 1$, estimando c_1 por mínimos cuadrados ordinarios en un paso preliminar. Sin embargo, por lo general se acepta que los retornos no son pronosticables, así que es raro ver

valores de L_y diferentes de cero. Por su parte, van der Sluis (1997) propone y estima un modelo en el que la log-varianza sigue un proceso ARMA(p,q), sin modificaciones en la ecuación de retornos. A este modelo lo denomina ASARMAV(p,q). De forma similar, el modelo permite adicionar saltos en la ecuación de retornos (ver por ejemplo, Chib, Nardari y Shephard (2002)), para modelar periodos prolongados de altos o bajos retornos.

Asumiendo que la innovación $\{\varepsilon_t\}$ es normal estándar, Davis y Mikosch (2009a) establecen que el comportamiento asintótico de las colas de la distribución de los retornos está dado por:

$$P(y_t > x) = \frac{1}{2}P(Y_t > 2 \log x), \quad (2.17)$$

donde $Y_t = \log y_t^2$ y

$$P(Y_t > z) \sim \frac{\sigma_h^2}{\sqrt{\pi}} \exp \left\{ -\frac{z^2}{2\sigma_h^2} + \frac{z \log z + (k-1)z - (k + \sigma_h^2) \log z}{\sigma_h^2} - \frac{\log^2 z}{2\sigma_h^2} - \frac{k^2}{2\sigma_h^2} + O\left(\frac{\log^2 z}{z}\right) \right\}, \quad (2.18)$$

a medida que $z \rightarrow \infty$, donde $k = \log(2/\sigma_h^2)$. Por ende, el modelo SV log-normal con innovación de los retornos Gaussiana tiene colas más pesadas que la normal, pero esto puede no ser suficiente para modelar las observaciones extremas de los retornos. Note que un retorno extremo implica una volatilidad condicional alta en el modelo con innovación normal. Asumiendo una distribución de colas pesadas para ε_t , las observaciones extremas pueden ser absorbidas por ε_t en lugar de la volatilidad (Jacquier, Polson y Rossi, 2004). Esto implica que el modelo con colas pesadas es más resistente a valores extremos, lo cual tiene implicaciones importantes para la estimación del modelo y para la estimación y pronóstico de la volatilidad. Como ha sido observado por Liesenfeld y Jung (2000), incluir una innovación de cola pesada influye en la estimación de los parámetros de la ecuación de log-varianza produciendo un parámetro de persistencia más alto y un parámetro de variabilidad menor. Estos cambios en la estimación de los parámetros ayudan a mejorar considerablemente los pronósticos de la volatilidad (Jacquier, Polson y Rossi, 2004). Otro argumento en contra del modelo con innovación normal ha sido sugerido por Liesenfeld y Jung (2000), quienes sostienen que este modelo no puede explicar simultáneamente la distribución leptocúrtica de los retornos y el decrecimiento lento de las funciones de autocorrelación de los

retornos al cuadrado. Además, empíricamente se ha encontrado que las estimaciones de modelos SV con colas pesadas ofrecen mejores ajustes a los datos de retornos (Gallant, Hsieh y Tauchen, 1997).

La función de distribución de una variable aleatoria A tiene colas pesadas si obedece la ley de potencias:

$$P(A > x) = x^{-\alpha}L(x), \quad (2.19)$$

donde $\alpha > 0$ y L es una función de variación lenta, que incluye constantes, logaritmos, logaritmos iterados y potencias de logaritmos (Davis y Mikosch, 2009b). La distribución de A se dice que tiene “variación regular con índice α ”. Para el modelo sin efecto leverage, Davis y Mikosch (2009a) establecen que la distribución de los retornos hereda el mismo comportamiento de las colas pesadas de ε_t :

$$P(y_t > x) \sim \mathbb{E}(\sigma_t^\alpha)P(\varepsilon_t > x) \quad \text{y} \quad P(y_t \leq -x) \sim \mathbb{E}(\sigma_t^\alpha)P(\varepsilon_t \leq -x),$$

a medida que $x \rightarrow \infty$, siempre que $\mathbb{E}(\sigma_t^{\alpha+\delta}) < \infty$ para algún $\delta > 0$. En el caso de varianza condicional log-normal, $\mathbb{E}(\sigma_t^{\alpha+\delta})$ es finito para todo $\delta > 0$, por ende basta con asignar una distribución de colas pesadas a ε_t para que los retornos también tengan colas pesadas. De forma interesante, para una base de datos de retornos de acciones de 20 países, Jondeau y Rockinger (2003) encuentran que las dos colas de la distribución, derecha e izquierda, tienen índices similares.

El modelo de colas pesadas más utilizado es el que asume una distribución t escalada para ε_t , definida como:

$$\varepsilon_t = \frac{\sqrt{\frac{v-2}{v}}\delta_t}{\sqrt{\frac{\chi_{t,v}^2}{v}}}, \quad (2.20)$$

donde $\delta_t \sim \text{iid}N(0, 1)$ y $\chi_{t,v}^2 \sim \chi_v^2$, con v pequeño (Kim, Shephard y Chib, 1998). Los grados de libertad (v) pueden ser impuestos o estimados con los datos. Un modelo con distribución t ha sido estimado, entre otros, por Jacquier, Polson y Rossi (2004) y Chib, Nardari y Shephard (2002) en el contexto de estimación Bayesiana, y por Jungbacker y Koopman (2009) en el contexto de estimación Monte Carlo basada en simulaciones. Otras opciones son la distribución Normal Inversa Gaussiana (NIG) (Barndorff-Nielsen (1997) y Andersson (2001)) y la distribución de error generalizada (GED), propuesta por Nelson (1991) para los modelos EGARCH y aplicada por Liesenfeld y Jung (2000) a los modelos SV. Recientemente, se han propuesto modelos SV

con dos factores en la ecuación de log-varianza, uno para explicar comportamientos de corto plazo y otro para las tendencias de largo plazo. Modelos de este tipo tienen colas pesadas y han sido estimados por Liesenfeld y Richard (2003) y Chernov *et ál.* (2003), entre otros. Estos últimos autores estiman su modelo mediante EMM.

Una alternativa no paramétrica para colas pesadas fue propuesta por Gallant, Hsieh y Tauchen (1997), quienes estiman un modelo SV que contiene una transformación *spline* cuadrática de la innovación de los retornos. El modelo está dado por:

$$\begin{aligned} y_t &= \sigma_t T(\varepsilon_t), \\ T(\varepsilon_t) &= b_0 + b_1 \varepsilon_t + b_2 \varepsilon_t^2 + b_3 I_+(\varepsilon_t) \varepsilon_t^2, \\ h_t &= \alpha(1 - \phi) + \phi h_{t-1} + \sigma_\eta \eta_t, \end{aligned}$$

donde se imponen las restricciones $\mathbb{E}[T(\varepsilon_t)] = 0$ y $\text{Var}[T(\varepsilon_t)] = 1$ a los coeficientes b_j , $\varepsilon_t \sim \text{iid}N(0, 1)$ y la variable $I_+(\varepsilon_t)$ es una variable indicadora que toma el valor de 1 si $\varepsilon_t > 0$ y de 0 si $\varepsilon_t < 0$, y tiene como objetivo capturar la asimetría en la distribución de los retornos. Esta alternativa hace más gruesas las colas y permite asimetría, pero a costa de introducir más parámetros al modelo.

Capítulo 3

El Método de Momentos Eficiente

3.1. Introducción

En el modelo de volatilidad estocástica, los retornos se observan pero no la volatilidad, debido a que esta se considera una variable aleatoria latente. Esto implica que no es posible obtener una expresión analítica para la función de verosimilitud en términos de variables observadas, y por ende no es posible estimar el modelo directamente por Máxima Verosimilitud. Los primeros métodos analíticos para la estimación de modelos SV fueron el Método Generalizado de Momentos (GMM) y el Método de Cuasi-Máxima Verosimilitud (QML) aplicado a una transformación del modelo. Luego aparecieron métodos que recurren a simulaciones.

El GMM, propuesto por Hansen (1982) y aplicado por Melino y Turnbull (1990) a la estimación de modelos de volatilidad estocástica, se basa en la idea de minimizar un vector de diferencias entre momentos poblacionales y momentos muestrales, multiplicado por una matriz de ponderaciones, que es la inversa de una estimación de la matriz de covarianzas asintótica de las condiciones de momentos. Mediante estudios de simulaciones, Andersen y Sørensen (1996), Chumacero (1997) y Jacquier, Polson y Rossi (1994) mostraron que GMM es ineficiente en muestras finitas, ya que los estimadores arrojan altos RMSE y tienen sesgos sustanciales. El GMM tiene problemas especialmente cuando el parámetro ϕ es cercano a uno (persistencia de la volatilidad) o cuando el coeficiente de variación al cuadrado es bajo, casos en los que

los algoritmos tienden a colapsar a soluciones de esquina. De hecho, Jacquier, Polson y Rossi (1994) reportan que para un tamaño de muestra de 500, cerca de la mitad de las estimaciones se fueron a soluciones de esquina. Andersen y Sørensen (1996) muestran que este problema se reduce a medida que aumenta el tamaño de muestra, sin embargo la convergencia de los estimadores es lenta, requiriéndose tamaños de muestra superiores a 4.000 para una eficiencia aceptable. Andersen y Sørensen (1996) sostienen que la deficiencia del GMM en muestras pequeñas se debe en gran medida a que estima mal la matriz de covarianzas asintótica. Además de lo anterior, GMM tiene el inconveniente que pueden presentarse óptimos globales por fuera de la región estacionaria o que la superficie criterio no sea convexa (Jacquier, Polson y Rossi, 1994).

El método QML fue propuesto por Harvey, Ruiz y Shephard (1994). La idea es que al tomar logaritmo al cuadrado de los retornos, el modelo SV se transforma en un modelo lineal de espacio estado, al cual se aplica el filtro de Kalman. El principal problema de este método es que la innovación resultante en el modelo transformado difícilmente sea Gaussiana, lo que causa que el método sea ineficiente en muestras pequeñas, aunque el estimador es consistente y asintóticamente normal (Kim, Shephard y Chib, 1998). Jacquier, Polson y Rossi (1994) encuentran que, al igual que GMM, QML es deficiente para datos con coeficiente de variación bajo y alta persistencia de la volatilidad. Un problema adicional del QML es que al tomar cuadrados a los retornos se pierde información, por ejemplo el efecto leverage (Ghysels, Harvey y Renault, 1996).

Estas deficiencias de los métodos analíticos llevaron al surgimiento de métodos que recurren o se basan en simulaciones (para una recopilación de los principales métodos, ver Broto y Ruiz (2004)). Uno de estos es el Método de Momentos Eficiente (EMM), propuesto por Gallant y Tauchen (1996), el cual ha sido utilizado para la estimación de modelos de volatilidad estocástica en tiempo discreto y en tiempo continuo, modelos de tasas de interés y tasas de cambio, entre otros (ver Gallant y Tauchen (2010a)). Básicamente, el EMM consiste en la aplicación de un criterio GMM en el que las condiciones de momentos están dadas por las esperanzas de los scores de un modelo auxiliar bajo el modelo estructural de volatilidad estocástica. Los datos entran a través de los estimadores de máxima verosimilitud de los parámetros del modelo auxiliar, los cuales se emplean para evaluar los scores. Debido a que en la mayoría de los casos las esperanzas no se pueden obtener analíticamente, estas se

deben aproximar a través de simulaciones.

Los pasos del EMM son los siguientes. En primer lugar, se resumen las características de los datos aplicando QML a una densidad de transición (modelo auxiliar), que debe ser una aproximación cercana al verdadero proceso generador de datos. Luego, se utilizan simulaciones para evaluar la esperanza de los scores auxiliares bajo el modelo de volatilidad estocástica, donde los parámetros auxiliares se reemplazan por sus estimadores QML del paso anterior. Finalmente, se optimiza una función criterio chi-cuadrado, como en GMM, para obtener los estimadores EMM de los parámetros del modelo SV.

Aparte de eliminar la arbitrariedad en la selección de los momentos de GMM, el EMM ha demostrado tener numerosas ventajas. Chumacero (1997) y Andersen, Chung y Sorensen (1999) comparan la capacidad de ambos métodos para estimar el modelo SV sin efecto leverage. En ambos estudios, el EMM es claramente superior a GMM en términos de eficiencia, dado que arroja menores RMSE, estimadores menos sesgados y menores tasas de falla en la convergencia. Otra ventaja del EMM es que los RMSE declinan exponencialmente con el tamaño de la muestra, mientras que el GMM no tiene esta característica (Chumacero, 1997). Por último, cuando el modelo auxiliar es una buena aproximación de la verdadera densidad condicional, en EMM no se requiere acudir a una complicada matriz de ponderaciones como en GMM. Andersen, Chung y Sorensen (1999) muestran que EMM también es superior a QML y es una alternativa competente a métodos Bayesianos computacionalmente más complejos de implementar.

En el resto del capítulo se explica el método EMM detalladamente, se resumen sus propiedades asintóticas y se destaca el papel del modelo auxiliar.

3.2. Descripción y Propiedades del EMM

Esta descripción del método se hace siguiendo a Gallant y Tauchen (2002, 2010a), Gallant, Hsieh y Tauchen (1997) y Andersen, Chung y Sørensen (1999).

Denote el vector de parámetros del modelo de volatilidad estocástica definido en las ecuaciones (1.1)-(1.2) y (1.5), o “modelo estructural”, por $\lambda = (\alpha, \phi, \sigma_\eta, \rho)'$, y sea $p(y_{t-L}, y_{t-L+1}, \dots, y_t | \lambda)$ la densidad estacionaria conjunta de los retornos $y_{t-L}, y_{t-L+1}, \dots, y_t$

para este modelo. En la sección 1.2 vimos que la condición $0 < \phi < 1$ garantiza que el modelo SV es estacionario estricto y ergódico.

Sean $\{\tilde{y}_t\}_{t=1}^n$ los datos observados, que se presume son generados por un proceso SV con un vector de parámetros $\lambda^0 \in \Lambda \subset \mathbb{R}^{\ell_\lambda}$, donde ℓ_λ es la longitud del vector λ^0 . Suponga que es posible encontrar un modelo auxiliar que se ajuste razonablemente bien a los datos observados. El modelo auxiliar tiene vector de parámetros $\theta \in \Theta \subset \mathbb{R}^{\ell_\theta}$, con $\ell_\theta > \ell_\lambda$, y densidad condicional

$$f(y_t|x_{t-1}, \theta),$$

donde $x_{t-1} = (y_{t-L}, y_{t-L+1}, \dots, y_{t-1})$ representa el vector de rezagos de los retornos.

El primer paso del EMM consiste en estimar θ por QML con los datos observados:

$$\tilde{\theta}_n = \arg \max_{\theta \in \Theta} \frac{1}{n} \sum_{t=1}^n \log f(\tilde{y}_t|\tilde{x}_{t-1}, \theta). \quad (3.1)$$

Si el modelo auxiliar es una buena aproximación a los datos, debe comportarse bien en pruebas de especificación estadística. Cabe destacar que, de hecho, esta es la única manera en que los datos reales ingresan al método. Ahora defina el criterio

$$m(\lambda, \theta) = \int \cdots \int \frac{\partial}{\partial \theta} \log f(y_t|x_{t-1}, \theta) p(y_{t-L}, \dots, y_t|\lambda) dy_{t-L} \cdots dy_t. \quad (3.2)$$

Es decir, $m(\lambda, \theta)$ es el valor esperado de los scores de la distribución condicional del modelo auxiliar bajo el modelo estructural de volatilidad estocástica. Como es indicado por Gallant, Hsieh y Tauchen (1997), el fundamento del EMM consiste en que el estimador QML del modelo auxiliar, $\tilde{\theta}_n$, induce el valor de θ para el cual $m(\lambda^0, \theta) = 0$. Por ende, $m(\lambda, \tilde{\theta}_n)$ debe estar cerca de cero para valores de λ cercanos a λ^0 . Luego, el objetivo de EMM es minimizar $m(\lambda, \tilde{\theta}_n)$. En este aspecto, el EMM guarda cierta similitud con Máxima Verosimilitud, dado que este último puede entenderse como un método de momentos en el que los momentos están dados por los scores de la función de verosimilitud.

Como $m(\lambda, \tilde{\theta}_n)$ en la mayoría de los casos no se puede obtener por solución cerrada, se debe aproximar por integración de Monte Carlo, promediando sobre una serie simulada del modelo SV, lo cual es posible porque el modelo es estacionario y ergódico:

$$m(\lambda, \tilde{\theta}_n) \approx \frac{1}{N} \sum_{\tau=1}^N \frac{\partial}{\partial \theta} \log f(\hat{y}_\tau(\lambda)|\hat{x}_{\tau-1}(\lambda), \tilde{\theta}_n), \quad (3.3)$$

donde $\{\hat{y}_\tau(\lambda)\}_{\tau=1}^N$ es una larga serie simulada del modelo SV para un vector de parámetros λ . En este aspecto, EMM es similar al método simulado de momentos (SMM) de Duffie y Singleton (1993), el cual obtiene los estimadores al minimizar un criterio GMM en el que los momentos poblacionales se obtienen, no por vía analítica, sino calculándolos con base en largas simulaciones del modelo a estimar. Se debe tener en cuenta que para calcular los valores de $m(\lambda, \tilde{\theta}_n)$, primero se debe correr la simulación hasta cierto número de observaciones con el fin de eliminar oscilaciones temporales ocasionadas por los valores iniciales. Las simulaciones introducen un error adicional, pero este se hace despreciable fijando un valor de N arbitrariamente grande.

Finalmente, el estimador EMM del vector de parámetros se obtiene al minimizar un criterio tipo GMM:

$$\tilde{\lambda}_n = \arg \min_{\lambda \in \Lambda} [m'(\lambda, \tilde{\theta}_n)(\tilde{j}_n)^{-1}m(\lambda, \tilde{\theta}_n)], \quad (3.4)$$

donde $(\tilde{j}_n)^{-1}$ es una matriz de ponderaciones. Si el modelo auxiliar es una buena aproximación del modelo estructural, Gallant y Tauchen (1996) establecen que una matriz de ponderaciones aceptable es la inversa del estimador de la matriz de covarianzas asintóticas del vector score, dada por:

$$\tilde{j}_n = \frac{1}{n} \sum_{t=1}^n \left[\frac{\partial}{\partial \theta} \log f(\tilde{y}_t | \tilde{x}_{t-1}, \tilde{\theta}_n) \right] \left[\frac{\partial}{\partial \theta} \log f(\tilde{y}_t | \tilde{x}_{t-1}, \tilde{\theta}_n) \right]'. \quad (3.5)$$

Note que la matriz \tilde{j}_n se puede calcular inmediatamente después de obtener los estimadores QML del vector de parámetros auxiliares. Gallant y Long (1997) demuestran que EMM es tan eficiente como máxima verosimilitud cuando el modelo auxiliar abarca todas las características del modelo estructural. Por lo tanto, esta es la cota de eficiencia del EMM. Sin embargo, la eficiencia máxima difícilmente se logre, pero si la densidad auxiliar es una buena aproximación de la verdadera distribución de los datos, la eficiencia del EMM es casi plena y la matriz de covarianzas asintótica de los estimadores se aproxima a la de máxima verosimilitud (Gallant y Tauchen (1999), Gallant y Long (1997)). Lo importante es que una adecuada elección del modelo auxiliar permite utilizar la matriz de ponderaciones indicada anteriormente, sin necesidad de recurrir a matrices más elaboradas, como ocurre en GMM (matrices de ponderación HAC). Si el modelo auxiliar no aproxima adecuadamente la verdadera densidad condicional, podrían aparecer problemas en la estimación, como sesgos o colapsos del algoritmo, y además se calcularían mal los errores estándar, lo cual conllevaría a inferencias erróneas acerca de los parámetros del modelo.

3.2.1. Propiedades asintóticas del estimador EMM y pruebas de especificación

Bajo condiciones de regularidad estándares (Gallant y Tauchen, 1996),

$$\hat{\lambda}_n \xrightarrow{a.s.} \lambda^0.$$

Es decir, el estimador EMM es consistente fuerte. Además, su distribución asintótica es normal:

$$\sqrt{n}(\hat{\lambda}_n - \lambda^0) \xrightarrow{d} N(0, [(M^0)'(j^0)^{-1}(M^0)]^{-1}), \quad (3.6)$$

donde $M^0 = M(\lambda^0, \theta^0)$ y

$$M(\lambda, \theta) = \frac{\partial m(\lambda, \theta)}{\partial \lambda'}. \quad (3.7)$$

M^0 puede ser estimado consistentemente por

$$\hat{M}_n = M_n(\hat{\lambda}_n, \tilde{\theta}_n) \xrightarrow{a.s.} M^0.$$

Por otro lado, $\tilde{j}_n \xrightarrow{a.s.} j^0$, donde

$$j^0 = \mathbb{E}_{\lambda^0} \left\{ \left[\frac{\partial}{\partial \theta} \log f(y_0 | x_{-1}, \theta^0) \right] \left[\frac{\partial}{\partial \theta} \log f(y_0 | x_{-1}, \theta^0) \right]' \right\}. \quad (3.8)$$

Intervalos de confianza de tipo Wald para los elementos λ_i del vector λ pueden ser construidos calculando \hat{M}_n numéricamente y obteniendo la varianza asintótica estimada de $\hat{\lambda}_i$ como el i -ésimo elemento de la diagonal de la matriz $(1/n)[(\hat{M}_n)'(\tilde{j}_n)^{-1}(\hat{M}_n)]^{-1}$.

Para chequear la especificación del modelo, bajo la hipótesis nula de que el modelo SV $p(y_{t-L}, \dots, y_t | \lambda)$ es correcto,

$$L_0 = n \cdot m'(\hat{\lambda}_n, \tilde{\theta}_n)(\tilde{j}_n)^{-1}m(\hat{\lambda}_n, \tilde{\theta}_n) \xrightarrow{d} \chi_{(\ell_\theta - \ell_\lambda)}^2, \quad (3.9)$$

donde ℓ_θ y ℓ_λ son las longitudes de los vectores θ y λ , respectivamente. Entonces, una condición necesaria para la identificación es $\ell_\theta > \ell_\lambda$, es decir, el modelo auxiliar debe tener más parámetros que el modelo de volatilidad estocástica. Las condiciones suficientes se discuten en Gallant y Tauchen (1996). Valores grandes del estadístico L_0 llevan a rechazar la hipótesis de que el modelo SV es adecuado para describir los datos.

Si se rechaza la hipótesis nula, se pueden inspeccionar los cuasi-t-ratios para tener un indicio de cuál es el motivo por el que el modelo SV no se ajusta bien a los datos:

$$\hat{T}_n = S_n^{-1} \sqrt{nm}(\hat{\lambda}_n, \tilde{\theta}_n); \quad S_n = \left[\text{diag} \left\{ \tilde{j}_n - (\hat{M}_n)[(\hat{M}_n)'(\tilde{j}_n)^{-1}(\hat{M}_n)]^{-1}(\hat{M}_n)' \right\} \right]^{\frac{1}{2}}. \quad (3.10)$$

Valores grandes de los cuasi-t-ratios revelan características de los datos que el modelo SV no puede aproximar (Gallant, Hsieh y Tauchen, 1997).

3.3. Síntesis

El método EMM se puede sintetizar en los siguientes pasos:

1. Elegir un modelo auxiliar con vector de parámetros θ , que se ajuste bien a los datos.
2. Obtener las expresiones analíticas de los scores del modelo auxiliar.
3. Estimar θ por QML (aplicar (3.1)).
4. Calcular la matriz \tilde{j}_n de (3.5) y obtener su inversa, la matriz de ponderaciones.
5. Programar la expresión (3.3), los valores esperados de los scores bajo una larga serie de datos simulados del modelo SV y evaluados en los parámetros auxiliares estimados por QML, $\tilde{\theta}_n$, obtenidos en el paso 3. Estas expresiones son los insumos para el criterio de optimización.
6. Aplicar el criterio (3.4) para obtener los estimadores EMM.
7. Una vez obtenidos los estimadores EMM, estos se introducen en las expresiones (3.3) y con estas se puede calcular el valor del estadístico (3.9), para chequear la hipótesis de que el modelo SV es correcto.
8. Las derivadas en (3.7) se pueden calcular numéricamente para obtener intervalos de confianza de tipo Wald, tal como se explica en la sección 3.2.1, y para obtener los valores de los cuasi-t-ratios dados en (3.10).

Capítulo 4

Simulaciones y Estimaciones

4.1. Introducción

En este capítulo se presentan las estimaciones por EMM de los modelos de volatilidad estocástica con datos simulados, recurriendo a varios modelos auxiliares y tamaños de muestra. Se busca, por un lado, determinar si el EMM es un buen método para estimar modelos SV con efecto leverage. Por otro lado, se pretende establecer el papel del modelo auxiliar en las estimaciones, comparando la eficiencia del EMM bajo modelos auxiliares GARCH, que no tienen leverage, y modelos EGARCH y APARCH, que efectivamente reproducen este efecto. Con el propósito de evaluar la capacidad de EMM para estimar un espectro más amplio de modelos de volatilidad estocástica, también se simula y estima un modelo SV con colas pesadas en la distribución de la innovación de la ecuación de retornos.

El capítulo está organizado de la siguiente manera: en la sección 4.2 se describen los modelos auxiliares que se van a utilizar en la estimación, se derivan los scores del modelo APARCH y se describe la metodología de simulación y estimación. Los resultados de las estimaciones para los modelos SV con innovación Gaussiana y con colas pesadas se presentan en la sección 4.3. Finalmente, en la sección 4.4 se hace una síntesis global de los resultados y se concluye bajo cuáles circunstancias el EMM es un método aceptable para estimar esta clase de modelos de volatilidad estocástica.

La estimación de modelos SV mediante EMM ha sido implementada en los programas S-Plus[®] (Zivot y Wang, 2006), SAS (Instituto SAS, en línea) y Ox[®] (van der Sluis, 1997), y también se encuentra disponible un código escrito en lenguaje C++, proveído por Gallant y Tauchen (2010a). En esta tesis, la estimación se hace con base en el lenguaje R (ver código en apéndice B).

4.2. Configuración de las Simulaciones

El primer modelo SV que se va a estimar es un modelo con distribución normal estándar de la innovación de los retornos, $\{\varepsilon_t\}$, y parámetros reales $(\alpha, \phi, \sigma_\eta, \rho)' = (-7.81, 0.95, 0.3273, -0.7)'$. Este modelo está calibrado para un coeficiente de variación al cuadrado de 2, y solo difiere ligeramente de los modelos considerados por Jacquier, Polson y Rossi (1994), Andersen, Chung y Sørensen (1999) y Giraldo y Lopera (2006). Se consideran tamaños de muestra de 500, 1.000, 2.000 y 4.000. Más adelante se estudia un modelo en el que $\{\varepsilon_t\}$ sigue una distribución t estandarizada de cola pesada.

Ahora se definen los modelos auxiliares a emplear en la estimación por EMM. De acuerdo con los resultados de la sección 1.2, los modelos que mejor reproducen la varianza condicional generada por un proceso SV log-normal son EGARCH y APARCH. El modelo GARCH se deja como referencia, para comparar la eficiencia de EMM bajo este modelo que no reproduce el efecto leverage. El GJR no se tiene en cuenta por ser un caso particular del modelo APARCH, además que como se apreció en la sección 1.2, éste tiende a sobrestimar considerablemente la verdadera varianza condicional.

Sin embargo, los modelos de heteroscedasticidad condicional Gaussianos no pueden, por sí solos, recoger todas las características de los retornos generados por los procesos de volatilidad estocástica (véase, por ejemplo, Carnero, Peña y Ruiz (2004)). Por esta razón, un modelo auxiliar adecuado requiere una densidad condicional que sea versátil para capturar características no Gaussianas de la distribución de los retornos. La densidad semiparamétrica (SNP), propuesta por Gallant y Nychka (1987) y Gallant y Tauchen (1989), reúne tales propiedades, por lo que se ha convertido en un estándar en la estimación por EMM.

Defina $z_t = (y_t - \mu_t)/\sigma_t$ como la innovación de los retornos $\{y_t\}$, donde σ_t es un parámetro o función de escala y μ_t es una función de localización, que se asume

depende linealmente de $x_{t-1} = (y_{t-1}, y_{t-2}, \dots)$, el vector de retornos pasados (o transformaciones de estos). La SNP es una distribución condicional para z_t de la forma:

$$h_k(z_t|x_{t-1}, \theta) = \frac{[P_k(z_t)]^2 \Phi(z_t)}{\int [P_k(u)]^2 \Phi(u) du}, \quad (4.1)$$

donde θ es el vector de parámetros de la SNP, $\Phi(z_t)$ es la densidad normal estándar y $P_k(z_t)$ es el polinomio:

$$P_k(z_t) = \sum_{i=0}^{K_z} b_i z_t^i, \quad b_0 = 1, \quad (4.2)$$

formado de las potencias de z_t , hasta de grado K_z . Note que si $K_z = 0$, se obtiene la distribución normal estándar. Para explicar la heteroscedasticidad condicional mediante la SNP, la función de escala σ_t puede ser modelada de forma paramétrica, acomodando una especificación de tipo GARCH. Una alternativa puramente no paramétrica consiste en permitir que los coeficientes del polinomio sean a su vez polinomios de grado K_x en x_{t-1} , formando el nuevo polinomio:

$$P_k^*(z_t, x_{t-1}) = \sum_{i=0}^{K_z} \left(\sum_{j=0}^{K_x} b_{ij} x_{t-1}^j \right) z_t^i, \quad b_{0,0} = 1. \quad (4.3)$$

Cuando $K_x > 0$, la forma de la distribución condicional $h_k(z_t|x_{t-1}, \theta)$ cambia con x_{t-1} , razón por la que este caso se denomina heterogéneo (Gallant, Hsieh y Tauchen, 1997). Sin embargo, esta especificación puramente no paramétrica de la densidad SNP tiene el problema que el índice K_x puede llegar a ser muy alto en las estimaciones para que el modelo pueda explicar la heteroscedasticidad presente en los datos, volviéndose una densidad no parsimoniosa difícil de estimar con muestras pequeñas. Como es reportado por Andersen, Chung y Sørensen (1999), esta versión de la SNP no arroja buenos resultados como modelo auxiliar en la estimación por EMM, siendo ampliamente superada por las versiones homogéneas ($K_x = 0$) con función de escala paramétrica. Por esta razón, en esta tesis solo se considera el caso homogéneo.

Haciendo el cambio de variable $y_t = \mu_t + \sigma_t z_t$, finalmente se llega a la distribución condicional de los retornos:

$$f_k(y_t|x_{t-1}, \theta) = \frac{1}{\sigma_t} \frac{[P_k((y_t - \mu_t)/\sigma_t)]^2 \Phi((y_t - \mu_t)/\sigma_t)}{\int [P_k(u)]^2 \Phi(u) du}. \quad (4.4)$$

La densidad SNP es un estimador no paramétrico consistente y eficiente de la verdadera densidad condicional (Gallant y Nychka (1987), Gallant y Long (1997)). Para la estimación por EMM, la SNP es preferible a los modelos con distribución condicional Gaussiana, puesto que agregando términos K_z , la densidad puede recoger características que son recurrentes en la distribución de los retornos, como colas pesadas y asimetría (Gallant y Nychka, 1987).

Con el objetivo de reducir la influencia de los valores extremos en las estimaciones de las densidades SNP, los cuales pueden inflar la log-verosimilitud, Zivot y Wang (2006, p. 902) sugieren aplicar una transformación spline al vector de rezagos x_{t-1} . Esta transformación está dada por:

$$\hat{x}_i = \begin{cases} 0.5[x_i - c - \ln(1 - x_i - c)], & x_i < -c \\ x_i, & -c < x_i < c \\ 0.5[x_i + c + \ln(1 + x_i - c)], & x_i > c. \end{cases} \quad (4.5)$$

En esta expresión, x_i representa los elementos del vector x_{t-1} . El valor de c se escoge arbitrariamente (para retornos diarios, puede estar entre 0.02 y 0.05). Esta transformación comprime los valores de x_i que exceden y están por debajo de c y $-c$, respectivamente. Estos rezagos transformados se utilizan para las estimaciones de las funciones de localización y escala.

Por otra parte, Gallant y Tauchen (2002, 2010b) argumentan que la SNP puede ser inestable cuando se utiliza como modelo auxiliar en la estimación por EMM, debido a que algunos valores suministrados dentro del proceso de simulación pueden ser muy pequeños. Para corregir este problema, agregan un término ϵ_0 a la SNP (sugieren un valor $\epsilon_0 = 0.001$), modificando la densidad condicional de la siguiente manera:

$$f_k^*(y_t|x_{t-1}, \theta) = \frac{1}{\sigma_t} \frac{\{[P_k((y_t - \mu_t)/\sigma_t)]^2 + \epsilon_0\} \Phi((y_t - \mu_t)/\sigma_t)}{\int [P_k(u)]^2 \Phi(u) du + \epsilon_0}. \quad (4.6)$$

Finalmente, el vector auxiliar θ contiene los parámetros de la ecuación de localización, los de la ecuación de escala y los coeficientes del polinomio de la densidad SNP, que dependen del valor del índice K_z . Para seleccionar el mejor modelo con datos reales, Gallant, Hsieh y Tauchen (1997) y Gallant y Tauchen (2010b) sugieren utilizar el criterio BIC.

Como se apreció en el capítulo 3, una de las condiciones del método EMM es que el número de elementos del vector de parámetros auxiliares debe ser mayor al

número de elementos del vector de parámetros del modelo SV. Al utilizar la densidad condicional SNP, se pueden obtener modelos auxiliares que cumplan este requisito sin necesidad de agregar parámetros a la parte de la función de escala, los cuales pueden generar problemas de estimación. Sin embargo, es necesario tener precaución con la sobreparametrización, especialmente para estimar modelos con muestras pequeñas. Andersen, Chung y Sørensen (1999) encuentran que para tamaños de muestra menores a 1.000, la estimación QML de modelos SNP tiende a llegar a óptimos locales. Por estas razones, estos autores recomiendan utilizar modelos condicionalmente normales o SNP con pocos parámetros cuando el tamaño de muestra es pequeño, mientras que para tamaños de muestra más grandes, agregar términos a la SNP mejora la eficiencia de los estimadores.

Algunos de los modelos auxiliares empleados en la literatura son densidades SNP con ecuaciones de escala que son versiones modificadas de los modelos ARCH y GARCH, en los que σ_t es una función de valores absolutos de rezagos de los retornos y de la propia σ_t , en contraste con las versiones clásicas que utilizan los retornos al cuadrado. Modelos de este tipo han sido empleados por Gallant, Hsieh y Tauchen (1997), Gallant y Tauchen (2002) y Zivot y Wang (2006, Cap. 23). Estos modelos auxiliares están implementados en el software S-Plus (Zivot y Wang, 2006, Cap. 23) y en el código en C++ proporcionado por Gallant y Tauchen (2010a). Gallant, Hsieh y Tauchen (1997), tratando de estimar un modelo SV con innovación spline, memoria larga y efecto leverage para el índice S&P500, utilizan dos modelos auxiliares. Un SNP con función de escala ARCH de 18 parámetros, con $K_z = 4$ y $K_x = 0$ y otro totalmente no paramétrico, con $K_z = 4$ y $K_x = 1$. Pero note que los modelos auxiliares no contienen efecto leverage ni memoria larga, lo cual puede explicar que el ARCH requiera 18 parámetros para intentar atrapar estas características. Por su parte, Andersen, Chung y Sørensen (1999) utilizan modelos SNP y Gaussianos con funciones de escala ARCH(q), GARCH(1,1) y EGARCH(1,1). Para la SNP utilizan $K_z = 2$ ó 4 con $K_x = 0$ y $K_z = 2$ con $K_x = 1$. Utilizando el modelo ARCH Gaussiano, encuentran problemas serios en la distribución de los estimadores EMM en muestras pequeñas, especialmente el parámetro ϕ , por lo que no recomiendan el uso de este modelo auxiliar. Giraldo y Lopera (2006) recurren a un SNP-EGARCH(1,1) con $K_z = 4$ para estimar modelos SV con efecto leverage. Chumacero (1997) utiliza un ARCH exponencial (logaritmo de la varianza en función de retornos cuadráticos), pero sin recurrir a la SNP. En lugar de ello, utiliza una distribución condicional t , para tratar

de atrapar las colas pesadas. En este estudio se utilizarán modelos auxiliares SNP homogéneos con funciones de escala GARCH, EGARCH y APARCH. Para tamaños de muestra pequeños, se consideran también los modelos condicionalmente Gaussianos.

4.2.1. Scores

En esta sección se presentan los scores del modelo auxiliar SNP-APARCH. Los scores del modelo SNP-GARCH son un caso particular de los del SNP-APARCH, mientras que los del SNP-EGARCH fueron obtenidos por Giraldo y Lopera (2006).

El modelo auxiliar SNP-APARCH tiene densidad condicional

$$f(y_t|x_{t-1}, \theta) = \frac{1}{\sqrt{2\pi}\sigma_t} \frac{[P^2(z_t) + \epsilon_0]e^{-\frac{z_t^2}{2}}}{\int_{-\infty}^{+\infty} P^2(u)\Phi(u)du + \epsilon_0}, \quad z_t = \frac{y_t - \mu_t}{\sigma_t} \quad (4.7)$$

formada con el polinomio

$$P(z_t) = 1 + \sum_{i=1}^{K_z} b_i z_t^i.$$

Dado que en general los retornos no tienen autocorrelaciones más allá del primer rezago, el factor de localización μ_t se asume que puede depender hasta de y_{t-1} :

$$\mu_t = a_0 + a_1 y_{t-1}. \quad (4.8)$$

Con respecto al factor de escala σ_t , se asume que sigue un proceso APARCH(1,1), dado por

$$\sigma_t^\delta = c_0 + c_1(|\varepsilon_{t-1}| - m_1 \varepsilon_{t-1})^\delta + d_1 \sigma_{t-1}^\delta, \quad (4.9)$$

donde $\varepsilon_{t-1} = y_{t-1} - \mu_{t-1}$. Se utiliza el modelo de orden (1,1), dado que es el caso más utilizado en la literatura para atrapar la heteroscedasticidad condicional.

El modelo SNP-APARCH tiene vector de coeficientes

$$\theta = (b_1, \dots, b_{K_z}, a_0, a_1, c_0, c_1, d_1, m_1, \delta)'$$

La log-verosimilitud de una observación y_t está dada por:

$$s(\theta) \propto -\frac{1}{2} \log \sigma_t^2 - \frac{z_t^2}{2} + \log[P^2(z_t) + \epsilon_0] - \log \left(\int_{-\infty}^{+\infty} P^2(u)\Phi(u)du + \epsilon_0 \right). \quad (4.10)$$

De la log-verosimilitud, es fácil ver que los scores de los coeficientes del polinomio b_1, \dots, b_{K_z} están dados por:

$$\frac{\partial s(\theta)}{\partial b_i} = \frac{2P(z_t)}{P^2(z_t) + \epsilon_0} \frac{\partial P(z_t)}{\partial b_i} - \frac{1}{\int_{-\infty}^{+\infty} P^2(u)\Phi(u)du + \epsilon_0} \frac{\partial}{\partial b_i} \left(\int_{-\infty}^{+\infty} P^2(u)\Phi(u)du \right). \quad (4.11)$$

donde

$$\frac{\partial P(z_t)}{\partial b_i} = z_t^i, \quad i = 1, \dots, K_z. \quad (4.12)$$

El denominador $\int_{-\infty}^{+\infty} P^2(u)\Phi(u)du + \epsilon_0$ y las derivadas $\frac{\partial}{\partial b_i}(\cdot)$ se obtienen al resolver la integral, que es una expresión que está dada en términos de los momentos hasta de orden $\mathbb{E}(u^{2K_z})$ de una variable $u \sim N(0, 1)$. Estos pueden calcularse fácilmente con la ayuda de algún software matemático.

Los scores de los parámetros de localización están dados por:

$$\frac{\partial s(\theta)}{\partial a_0} = \frac{1}{\sigma_t} \left[z_t - \frac{2P(z_t)}{P^2(z_t) + \epsilon_0} \frac{\partial P(z_t)}{\partial z_t} \right] + \frac{1}{2\sigma_t^2} \left[z_t^2 - 1 - \frac{2z_t P(z_t)}{P^2(z_t) + \epsilon_0} \frac{\partial P(z_t)}{\partial z_t} \right] \frac{\partial \sigma_t^2}{\partial a_0}, \quad (4.13)$$

y

$$\frac{\partial s(\theta)}{\partial a_1} = \frac{y_{t-1}}{\sigma_t} \left[z_t - \frac{2P(z_t)}{P^2(z_t) + \epsilon_0} \frac{\partial P(z_t)}{\partial z_t} \right] + \frac{1}{2\sigma_t^2} \left[z_t^2 - 1 - \frac{2z_t P(z_t)}{P^2(z_t) + \epsilon_0} \frac{\partial P(z_t)}{\partial z_t} \right] \frac{\partial \sigma_t^2}{\partial a_1}, \quad (4.14)$$

donde

$$\frac{\partial P(z_t)}{\partial z_t} = \sum_{j=1}^{K_z} j b_j z_t^{j-1}. \quad (4.15)$$

Por su parte, las derivadas $\frac{\partial \sigma_t^2}{\partial a_0}$ y $\frac{\partial \sigma_t^2}{\partial a_1}$ obedecen las ecuaciones recursivas:

$$\frac{\partial \sigma_t^2}{\partial a_0} = 2c_1 \sigma_t^{2-\delta} (|\varepsilon_{t-1}| - m_1 \varepsilon_{t-1})^{\delta-1} [I_{(-)}(z_{t-1} > 0) + m_1] + d_1 \sigma_t^{2-\delta} \sigma_{t-1}^{\delta-2} \frac{\partial \sigma_{t-1}^2}{\partial a_0}, \quad (4.16)$$

y

$$\frac{\partial \sigma_t^2}{\partial a_1} = 2c_1 \sigma_t^{2-\delta} (|\varepsilon_{t-1}| - m_1 \varepsilon_{t-1})^{\delta-1} [I_{(-)}(z_{t-1} > 0) + m_1] y_{t-2} + d_1 \sigma_t^{2-\delta} \sigma_{t-1}^{\delta-2} \frac{\partial \sigma_{t-1}^2}{\partial a_1}, \quad (4.17)$$

donde

$$I_{(-)}(z_{t-1} > 0) = \begin{cases} -1 & \text{si } \varepsilon_{t-1} > 0 \\ 1 & \text{si } \varepsilon_{t-1} < 0. \end{cases} \quad (4.18)$$

Esta variable indicadora surge al tomar la derivada del valor absoluto $|\varepsilon_{t-1}|$, la cual no está definida para $\varepsilon_{t-1} = 0$. Para no afectar el cálculo de los scores, si llegara a ocurrir algún $\varepsilon_{t-1} = 0$, lo cual es poco probable, éste se corrige reemplazándolo por un valor cercano a cero en la fórmula de la derivada.

Finalmente, los scores de los coeficientes de la ecuación de escala APARCH, $(c_0, c_1, d_1, m_1, \delta)$, están dados por:

$$\frac{\partial s(\theta)}{\partial c_0, c_1, d_1, m_1, \delta} = \frac{1}{2\sigma_t^2} \left(z_t^2 - \frac{2z_t P(z_t)}{P^2(z_t) + \epsilon_0} \frac{\partial P(z_t)}{\partial z_t} - 1 \right) \frac{\partial \sigma_t^2}{\partial c_0, c_1, d_1, m_1, \delta}, \quad (4.19)$$

donde las derivadas de σ_t^2 con respecto a los parámetros obedecen las siguientes ecuaciones recursivas:

$$\frac{\partial \sigma_t^2}{\partial c_0} = \frac{2\sigma_t^{2-\delta}}{\delta} \left(1 + \frac{d_1 \delta \sigma_{t-1}^{\delta-2}}{2} \frac{\partial \sigma_{t-1}^2}{\partial c_0} \right), \quad (4.20)$$

$$\frac{\partial \sigma_t^2}{\partial c_1} = \frac{2\sigma_t^{2-\delta}}{\delta} \left[(|\varepsilon_{t-1}| - m_1 \varepsilon_{t-1})^\delta + \frac{d_1 \delta \sigma_{t-1}^{\delta-2}}{2} \frac{\partial \sigma_{t-1}^2}{\partial c_1} \right], \quad (4.21)$$

$$\frac{\partial \sigma_t^2}{\partial d_1} = \frac{2\sigma_t^{2-\delta}}{\delta} \left(\sigma_{t-1}^\delta + \frac{d_1 \delta \sigma_{t-1}^{\delta-2}}{2} \frac{\partial \sigma_{t-1}^2}{\partial d_1} \right), \quad (4.22)$$

$$\frac{\partial \sigma_t^2}{\partial m_1} = \frac{2\sigma_t^{2-\delta}}{\delta} \left[-\delta c_1 \varepsilon_{t-1} (|\varepsilon_{t-1}| - m_1 \varepsilon_{t-1})^{\delta-1} + \frac{d_1 \delta \sigma_{t-1}^{\delta-2}}{2} \frac{\partial \sigma_{t-1}^2}{\partial m_1} \right], \quad (4.23)$$

$$\begin{aligned} \frac{\partial \sigma_t^2}{\partial \delta} &= \frac{2\sigma_t^{2-\delta}}{\delta} \left[c_1 (|\varepsilon_{t-1}| - m_1 \varepsilon_{t-1})^\delta \ln(|\varepsilon_{t-1}| - m_1 \varepsilon_{t-1}) + \frac{d_1 \sigma_{t-1}^\delta \ln \sigma_{t-1}^2}{2} \right] \\ &\quad - \frac{2\sigma_t^2 \ln \sigma_t^\delta}{\delta^2} + d_1 \sigma_t^{2-\delta} \sigma_{t-1}^{\delta-2} \frac{\partial \sigma_{t-1}^2}{\partial \delta}. \end{aligned} \quad (4.24)$$

4.2.2. Procedimiento de estimación

Se propone el siguiente esquema para la simulación y estimación de los modelos SV, para tamaños de muestra $n=500, 1.000, 2.000$ y 4.000 . Por cada tamaño de muestra, se hacen los siguientes pasos:

1. Se genera una serie simulada de los retornos de tamaño n del modelo SV respectivo. Para eliminar oscilaciones causadas por los valores iniciales, se genera y descarta una premuestra de 1.000 observaciones.

2. Para la estimación de los modelos auxiliares, se procede de la siguiente manera. Siguiendo las recomendaciones de Zivot y Wang (2006, cap. 23), se aplica la transformación spline (4.5) a los rezagos de los retornos, con un valor de c de 0.03. Con estos datos se estiman los parámetros de las ecuaciones de localización, de escala para especificaciones GARCH, EGARCH y APARCH y los coeficientes del polinomio de la densidad SNP. En la estimación del modelo auxiliar, se parte de diferentes conjuntos de valores iniciales, con el objetivo de evitar caer en óptimos locales. Este es un paso crítico, puesto que en las estimaciones se encontró que un óptimo local del modelo auxiliar puede llevar a una solución de esquina en la estimación EMM. Luego de estimar los modelos auxiliares, se recogen los parámetros estimados por QML en el vector $\tilde{\theta}_n$. Esto completa el primer paso del EMM.
3. Se evalúan los scores en los parámetros estimados y se obtiene la matriz \tilde{j}_n , dada en la ecuación (3.5).
4. Se aplica un optimizador a la función criterio (3.4), para obtener los estimadores EMM. El optimizador se programa generando simulaciones de tamaño 80.000 del modelo SV respectivo. En este paso se detectó una alta sensibilidad de la función criterio EMM a los valores iniciales del algoritmo. Cuando se escoge un vector de valores iniciales lejano a los verdaderos parámetros, las estimaciones tienden a ser considerablemente sesgadas. Por esta razón, se optó por iniciar el optimizador en el vector de parámetros verdaderos, tal como hicieron Andersen, Chung y Sørensen (1999) en su estudio Monte Carlo. No obstante, este hecho sugiere que al trabajar con datos reales, de los cuales se desconoce el verdadero proceso generador de datos, se debe tener la precaución de efectuar la estimación EMM partiendo de diferentes puntos de la región criterio.
5. Se repiten los pasos 1 a 4, 500 veces, para obtener las series de parámetros estimados.

4.3. Resultados

4.3.1. Modelo SV log-normal

La selección de los modelos auxiliares para cada tamaño de muestra se hizo considerando modelos con dependencia hasta del primer rezago en la función de localización, mientras que para la función de escala se estimaron modelos de orden (1,1). Para el polinomio de la SNP se ensayaron valores de K_z entre 0 y 8. Se hicieron las estimaciones para algunas series simuladas de modelos SV para cada tamaño de muestra y se realizó una selección previa de modelos auxiliares, de acuerdo al criterio BIC y a la significancia de los parámetros. En este punto, se hizo evidente que modelos SNP con muchos parámetros no permiten, en general, evaluar la significancia de los parámetros, debido a problemas de invertibilidad de la matriz Hessiana.

Tamaño de muestra			
500	1.000	2.000	4.000
AR(1)-GARCH-SNP(2)	AR(1)-GARCH-SNP(4)	AR(1)-GARCH-SNP(4)	AR(1)-GARCH-SNP(4)
EGARCH	EGARCH-SNP(4)	EGARCH-SNP(4)	EGARCH-SNP(4)
APARCH	APARCH-SNP(4)	APARCH-SNP(4)	APARCH-SNP(4)

Tabla 4.1: Modelos auxiliares empleados en la estimación por EMM del modelo de volatilidad estocástica. En todos los casos, los modelos de heteroscedasticidad condicional son de orden (1,1).

Enseguida, se procedió a realizar un experimento preliminar de estimaciones EMM utilizando los modelos auxiliares preseleccionados. Los modelos auxiliares finales se escogieron teniendo en cuenta los RMSE, la convergencia y la distribución de los estimadores de este estudio preliminar, descartando aquellos que producen altos RMSE, estimadores altamente sesgados, malas distribuciones de los estimadores y cantidades altas de estimaciones en valores de esquina. Los modelos auxiliares seleccionados se muestran en la tabla 4.1. Para el tamaño de muestra 500, se escogieron los modelos más parsimoniosos EGARCH(1,1) Gaussiano y APARCH(1,1) Gaussiano y un AR(1)-GARCH(1,1)-SNP(2). En este último caso, se hizo evidente que por no tener en cuenta el efecto leverage, fue necesario agregar términos SNP para mejorar las

estimaciones bajo el modelo auxiliar GARCH. Para los tamaños de muestra más grandes, se escogieron modelos EGARCH(1,1)-SNP(4), APARCH(1,1)-SNP(4) y AR(1)-GARCH(1,1)-SNP(4). En la escogencia de los modelos auxiliares, se buscó cierto equilibrio entre RMSE, convergencia del algoritmo y distribución de los estimadores. Agregando más términos SNP se reducen los RMSE hasta cierto punto, pero el porcentaje de estimaciones que caen en soluciones de esquina aumenta y las distribuciones de los estimadores resultan lejos de la normalidad, por lo cual esos modelos con más términos son menos atractivos.

Los principales resultados del experimento de Monte Carlo se presentan en la tabla 4.2. Por cada combinación de tamaño de muestra y modelo auxiliar, se reportan los promedios de los 500 parámetros estimados, los RMSE, desviaciones estándar, los valores P de la prueba de normalidad de Shapiro-Wilk y el porcentaje de estimaciones que caen en soluciones de esquina (colapsos), para los cuatro parámetros del modelo SV. En este punto vale la pena mencionar que durante el proceso de evaluación de los scores auxiliares con las series simuladas de modelos SV, la reconstrucción de la varianza en los modelos con términos EGARCH se fue a infinito en algunos casos, lo cual es causado por las características de la ecuación de recurrencia del EGARCH señaladas en la sección 1.2. Sin embargo, este problema se resolvió al reestimar utilizando una nueva serie de números aleatorios.

En lo que respecta a la convergencia del EMM, se observa una alta sensibilidad a la escogencia del modelo auxiliar. Las estimaciones basadas en modelos auxiliares con términos GARCH sufrieron de altas tasas de falla en la convergencia en los tamaños de muestra 500, 1.000 e, inclusive, 2.000. Se observa como al pasar de un modelo AR(1)-GARCH(1,1)-SNP(2) en el tamaño de muestra 500 a un modelo AR(1)-GARCH(1,1)-SNP(4) en el tamaño de muestra 1.000, el porcentaje de estimaciones que cayeron en soluciones de esquina sube de 21.8% a 28.3%. Por el contrario, al incluir términos SNP en los modelos EGARCH y APARCH, el porcentaje de colapsos cae. Estos últimos modelos mantienen tasas de colapsos bajas y similares y, de hecho, para el tamaño de muestra 4.000, ya ninguna estimación cayó en soluciones de esquina. Lo que esta evidencia sugiere es que las estimaciones basadas en modelos auxiliares que no proveen suficiente información acerca de las características de los datos, tenderán a generar colapsos en el criterio de optimización EMM.

Para el tamaño de muestra 500, el RMSE de $\hat{\alpha}$ bajo el modelo auxiliar SNP-GARCH es ligeramente inferior en comparación con los RMSE de los modelos au-

xiliares EGARCH y APARCH Gaussianos, pero sucede lo contrario con los demás parámetros. A partir del tamaño de muestra 1.000, el modelo SNP-GARCH tiene los RMSE más altos en todos los parámetros. Llama la atención principalmente el caso de $\hat{\rho}$, para el cual los RMSE de las estimaciones bajo este modelo auxiliar son considerablemente superiores a los RMSE bajo los modelos SNP-EGARCH y SNP-APARCH. Además, es notable que solo para este parámetro y modelo auxiliar, el RMSE no cae a medida que aumenta el tamaño de la muestra. Estos hechos ratifican la presunción de que una estimación basada en un modelo auxiliar que no reproduce el efecto leverage sería menos precisa en comparación con utilizar un modelo auxiliar que tuviera en cuenta este efecto.

Cuando se comparan los RMSE de los modelos auxiliares SNP-EGARCH y SNP-APARCH no se aprecian diferencias considerables, si bien en la mayoría de los casos los del modelo auxiliar SNP-APARCH son levemente inferiores. También se aprecia que a medida que aumenta el tamaño de muestra, los RMSE de todos los parámetros estimados bajo ambos modelos auxiliares se reducen y se van haciendo cada vez más similares. Por ende, los modelos auxiliares SNP-EGARCH y SNP-APARCH son competentes en términos de RMSE. Un análisis de la distribución de los estimadores, el sesgo y la dispersión permitirá establecer si existe alguna diferencia sustancial en las estimaciones bajo cada uno de estos modelos.

Tam. muestra	500				1.000			
Mod. Auxiliar	AR(1)-GARCH(1,1)-SNP(2)				AR(1)-GARCH(1,1)-SNP(4)			
Coefficientes	α	ϕ	σ_η	ρ	α	ϕ	σ_η	ρ
Promedio	-7.886	0.934	0.344	-0.684	-7.859	0.939	0.342	-0.677
RMSE	0.251	0.034	0.074	0.143	0.214	0.026	0.054	0.190
Desv. Estándar	0.239	0.030	0.072	0.142	0.209	0.024	0.052	0.188
Valor P test S-W	0.003	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Colapsos	21.8 %				28.3 %			
Mod. Auxiliar	EGARCH(1,1)				EGARCH(1,1)-SNP(4)			
Coefficientes	α	ϕ	σ_η	ρ	α	ϕ	σ_η	ρ
Promedio	-7.790	0.944	0.334	-0.712	-7.845	0.943	0.329	-0.710
RMSE	0.269	0.021	0.068	0.110	0.182	0.021	0.044	0.074
Desv. Estándar	0.269	0.020	0.068	0.110	0.179	0.020	0.044	0.074
Valor P test S-W	0.001	<0.001	<0.001	<0.001	0.533	<0.001	0.054	<0.001
Colapsos	5.5 %				3.9 %			
Mod. Auxiliar	APARCH(1,1)				APARCH(1,1)-SNP(4)			
Coefficientes	α	ϕ	σ_η	ρ	α	ϕ	σ_η	ρ
Promedio	-7.834	0.944	0.337	-0.715	-7.857	0.946	0.325	-0.722
RMSE	0.270	0.024	0.066	0.102	0.192	0.017	0.042	0.078
Desv. Estándar	0.269	0.023	0.065	0.101	0.186	0.016	0.042	0.075
Valor P test S-W	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.002	<0.001
Colapsos	5.1 %				1.4 %			
Tam. muestra	2.000				4.000			
Mod. Auxiliar	AR(1)-GARCH(1,1)-SNP(4)				AR(1)-GARCH(1,1)-SNP(4)			
Coefficientes	α	ϕ	σ_η	ρ	α	ϕ	σ_η	ρ
Promedio	-7.851	0.944	0.336	-0.679	-7.824	0.948	0.329	-0.694
RMSE	0.151	0.016	0.039	0.177	0.109	0.012	0.027	0.146
Desv. Estándar	0.145	0.014	0.038	0.176	0.108	0.012	0.027	0.146
Valor P test S-W	0.576	<0.001	0.020	<0.001	0.813	<0.001	<0.001	<0.001
Colapsos	10.2 %				3.5 %			
Mod. Auxiliar	EGARCH(1,1)-SNP(4)				EGARCH(1,1)-SNP(4)			
Coefficientes	α	ϕ	σ_η	ρ	α	ϕ	σ_η	ρ
Promedio	-7.836	0.946	0.328	-0.706	-7.828	0.948	0.329	-0.701
RMSE	0.139	0.015	0.032	0.051	0.092	0.008	0.023	0.037
Desv. Estándar	0.136	0.014	0.032	0.051	0.091	0.008	0.023	0.037
Valor P test S-W	0.996	<0.001	<0.001	0.114	0.489	<0.001	<0.001	0.986
Colapsos	0.6 %				0.0 %			
Mod. Auxiliar	APARCH(1,1)-SNP(4)				APARCH(1,1)-SNP(4)			
Coefficientes	α	ϕ	σ_η	ρ	α	ϕ	σ_η	ρ
Promedio	-7.843	0.947	0.328	-0.709	-7.822	0.949	0.326	-0.705
RMSE	0.131	0.012	0.030	0.056	0.087	0.007	0.021	0.039
Desv. Estándar	0.127	0.012	0.030	0.056	0.086	0.007	0.021	0.038
Valor P test S-W	0.279	<0.001	0.022	0.004	0.420	<0.001	0.593	0.586
Colapsos	0.6 %				0.0 %			

Tabla 4.2: Resultados de las estimaciones EMM del modelo de volatilidad estocástica log-normal (ecuaciones (1.1) y (1.6)), con $\varepsilon_t \sim N(0, 1)$, para tamaños de muestra de 500, 1.000, 2.000 y 4.000 y diferentes modelos auxiliares (ver tabla 4.1). Los parámetros verdaderos son $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$. La cuarta fila en cada bloque corresponde al valor P del test de normalidad de Shapiro-Wilk. Los resultados están basados en 500 estimaciones convergentes.

En las figuras 4.1 y 4.2 se presentan los boxplots de las estimaciones para los tamaños de muestra de 500 y 4.000, respectivamente. En el apéndice, figuras A1 y A2, se presentan las gráficas correspondientes a los tamaños de muestra 1.000 y 2.000. Para el tamaño de muestra 500, se observa que, si bien el estimador EMM de α bajo el modelo SNP-GARCH es el menos disperso, a la vez es el más sesgado. En cuanto al parámetro ϕ , el modelo SNP-EGARCH arroja los estimadores menos dispersos y el SNP-GARCH tiene el peor desempeño, arrojando los estimadores más sesgados, dispersos y con una cantidad considerable de valores extremos, todos ellos por debajo del valor real. Esto se repite para el caso de σ_η , en el cual también se observa que los modelos SNP-EGARCH y SNP-APARCH arrojan resultados similares. En lo que respecta a ρ , las estimaciones basadas en los modelos auxiliares SNP-EGARCH y SNP-APARCH son ligeramente sesgadas hacia abajo, mientras que el modelo auxiliar SNP-GARCH arroja los estimadores menos sesgados, pero con mayor dispersión.

Para el tamaño de muestra 4.000, presentado en la figura 4.2, ya se aprecia un deterioro considerable en las estimaciones basadas en el modelo auxiliar SNP-GARCH, puesto que este modelo arroja las estimaciones más dispersas y con mayor número de valores extremos para todos los parámetros, pero con especial énfasis en el parámetro de asimetría, ρ . En esta figura, también se observa que el modelo auxiliar SNP-APARCH arroja los mejores estimadores de los parámetros ϕ y σ_η , en cuanto a que presentan menos valores extremos que los estimadores EMM basados en SNP-EGARCH.

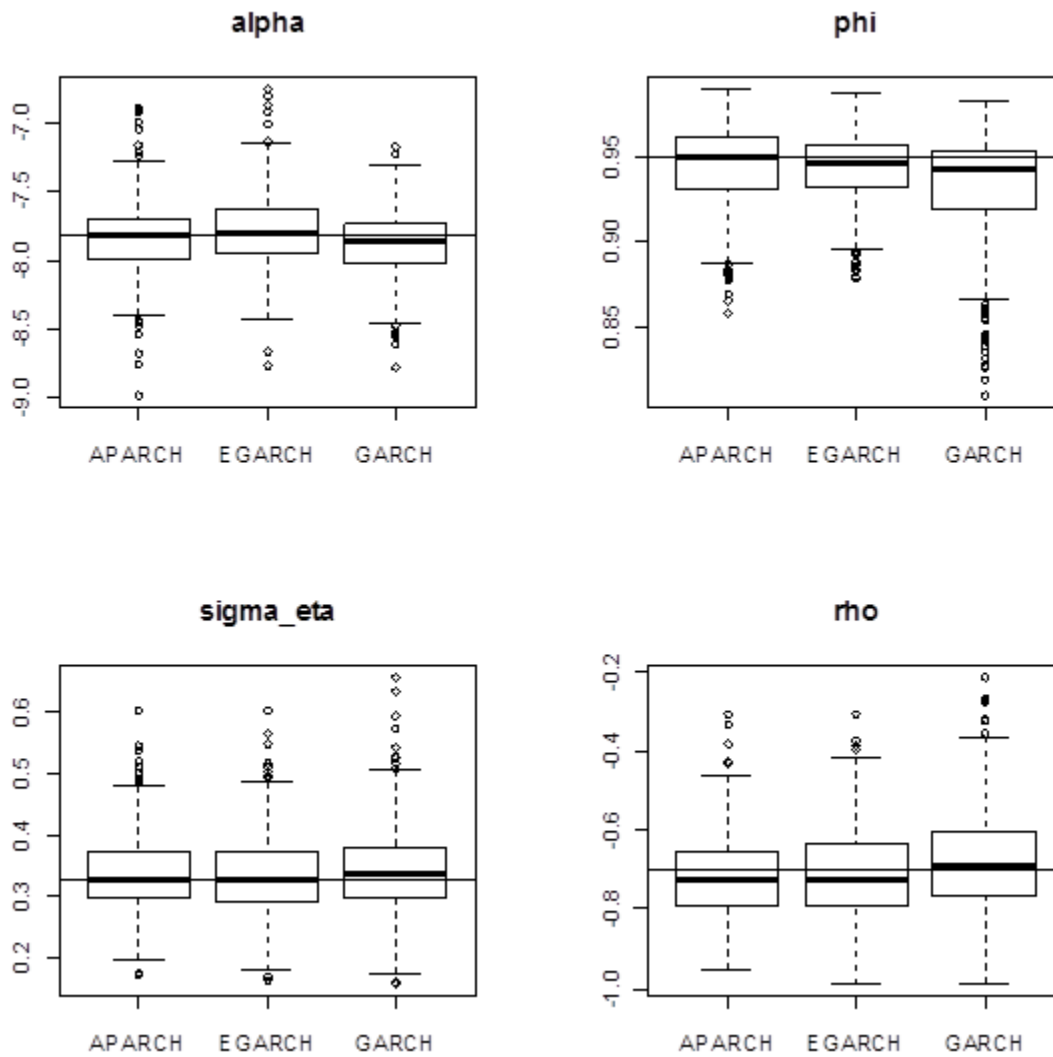


Figura 4.1: Resultados de las estimaciones EMM para el tamaño de muestra 500. Estos resultados están basados en 500 estimaciones convergentes dentro del espacio paramétrico. Las líneas horizontales corresponden a los verdaderos valores de los parámetros, que son $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

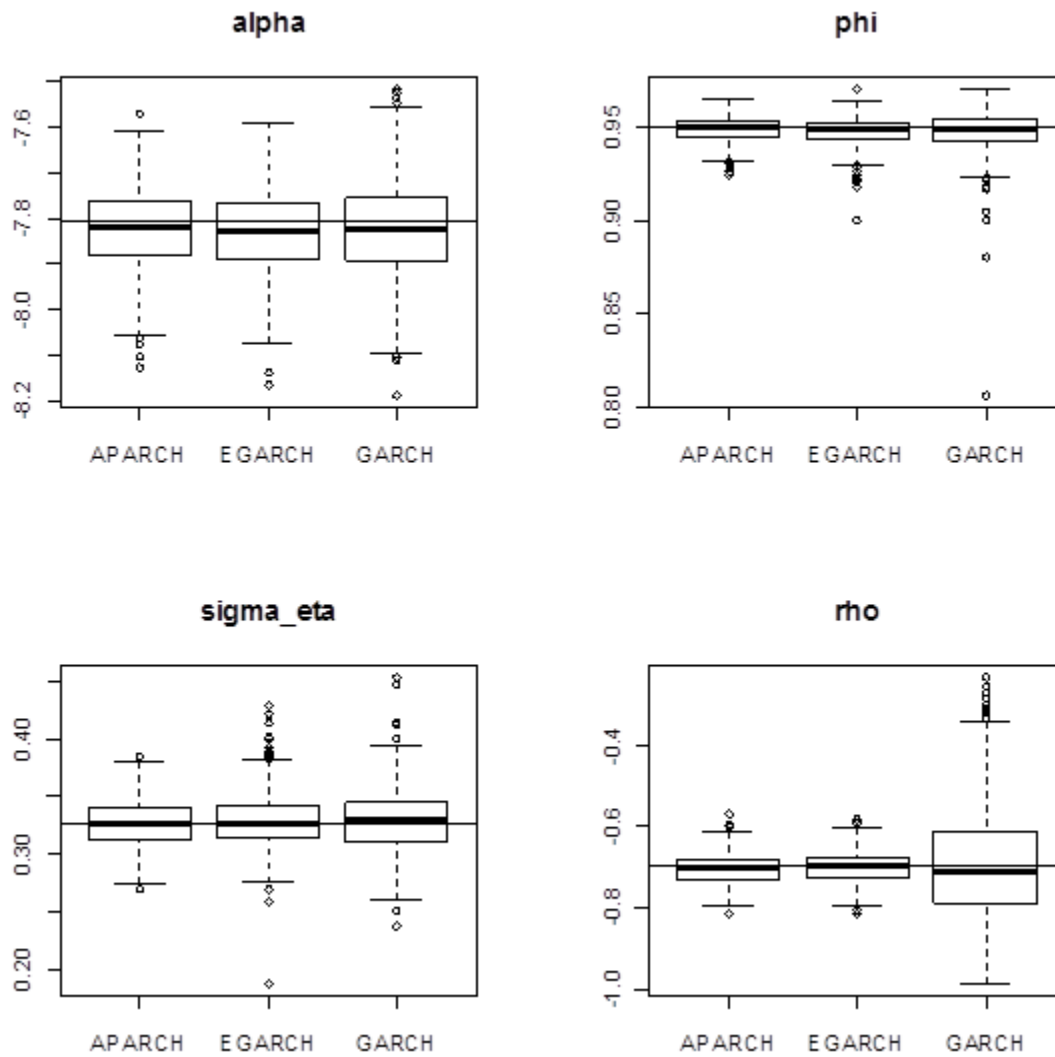


Figura 4.2: Resultados de las estimaciones EMM para el tamaño de muestra 4.000. Estos resultados están basados en 500 estimaciones convergentes dentro del espacio paramétrico. Las líneas horizontales corresponden a los verdaderos valores de los parámetros, que son $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

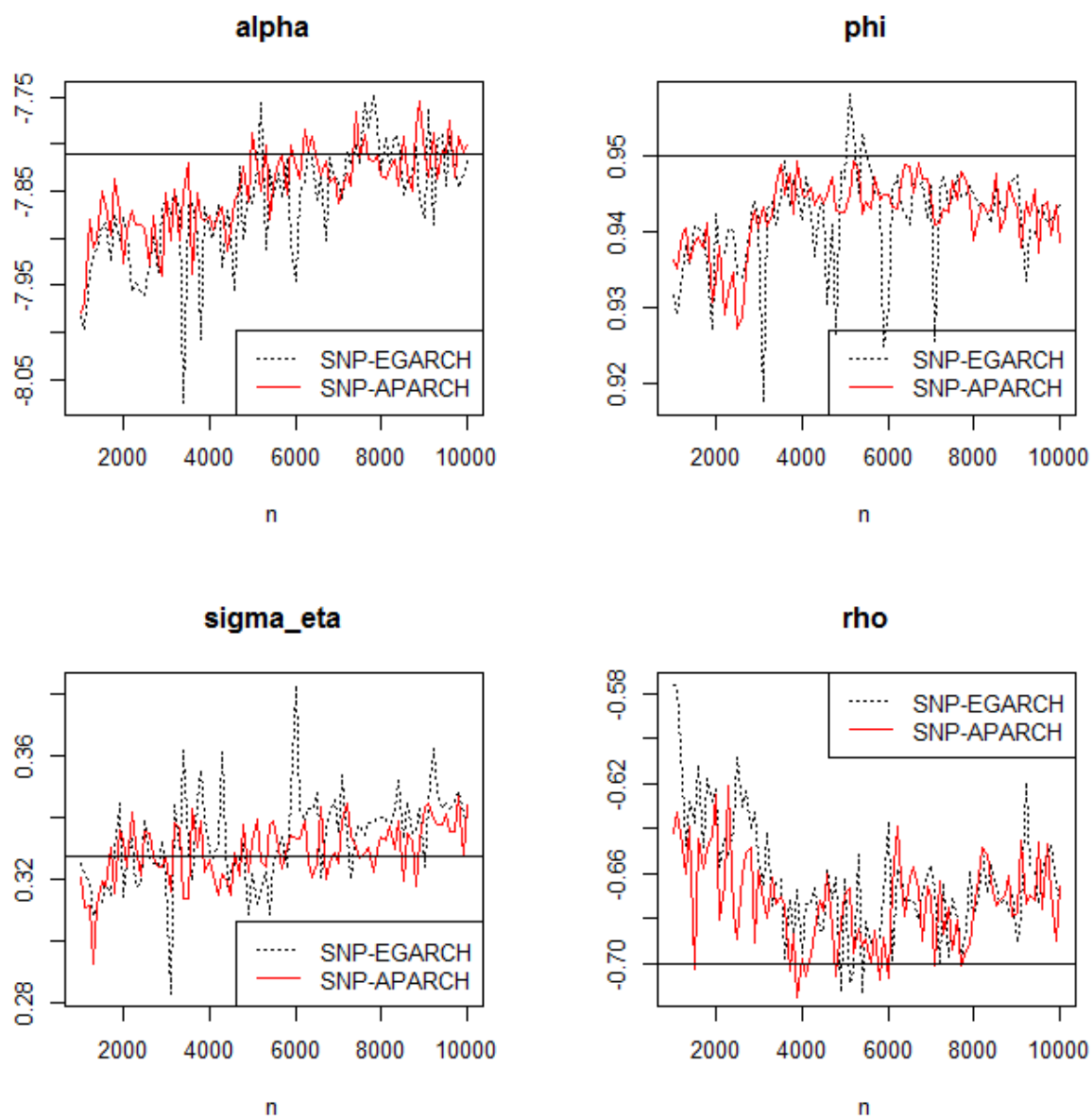


Figura 4.3: Comportamiento de los estimadores EMM a medida que aumenta el tamaño de muestra. Se comienza con un tamaño de muestra de 500 y se va incrementando de a 100 hasta llegar a 10.000. Los modelos auxiliares empleados en las estimaciones EMM son EGARCH(1,1)-SNP(4) y APARCH(1,1)-SNP(4). Las líneas horizontales corresponden a los verdaderos valores de los parámetros, que son $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

Anteriormente, se analizó el comportamiento de EMM sobre el promedio de 500 estimaciones para un mismo tamaño de muestra. En la figura 4.3, se evalúa cómo se comportan los estimadores al aumentar gradualmente el tamaño de muestra. Se empieza con un tamaño de muestra de 1.000 y se va aumentando de a 100 hasta alcanzar un tamaño de muestra de 10.000. En este experimento, se comparan las estimaciones bajo los modelos auxiliares EGARCH(1,1)-SNP(4) y APARCH(1,1)-SNP(4). Como se observa, las estimaciones mejoran rápidamente dentro de las primeras 3.000 observaciones, y luego se mantienen cerca de los verdaderos parámetros, especialmente en los casos de α y σ_η . Las estimaciones de ϕ y ρ tienden a desviarse en los tamaños de muestra más grandes, pero esto se atribuye a que se está manteniendo fijo el número de términos de la expansión SNP, cuando en realidad estos deben aumentar para que la densidad SNP abarque bien las principales características de los datos a medida que crece el tamaño de muestra (Gallant, Hsieh y Tauchen, 1997). En la figura también se aprecia que el parámetro en el que las estimaciones tienen una menor variabilidad de muestra a muestra es ϕ , mientras que las estimaciones de los demás parámetros se ven más afectadas al cambiar el tamaño de la muestra. Pero lo más importante es que en todos los casos se aprecia una mayor variabilidad de las estimaciones basadas en el modelo SNP-EGARCH, lo cual conforma un punto a favor del modelo auxiliar SNP-APARCH.

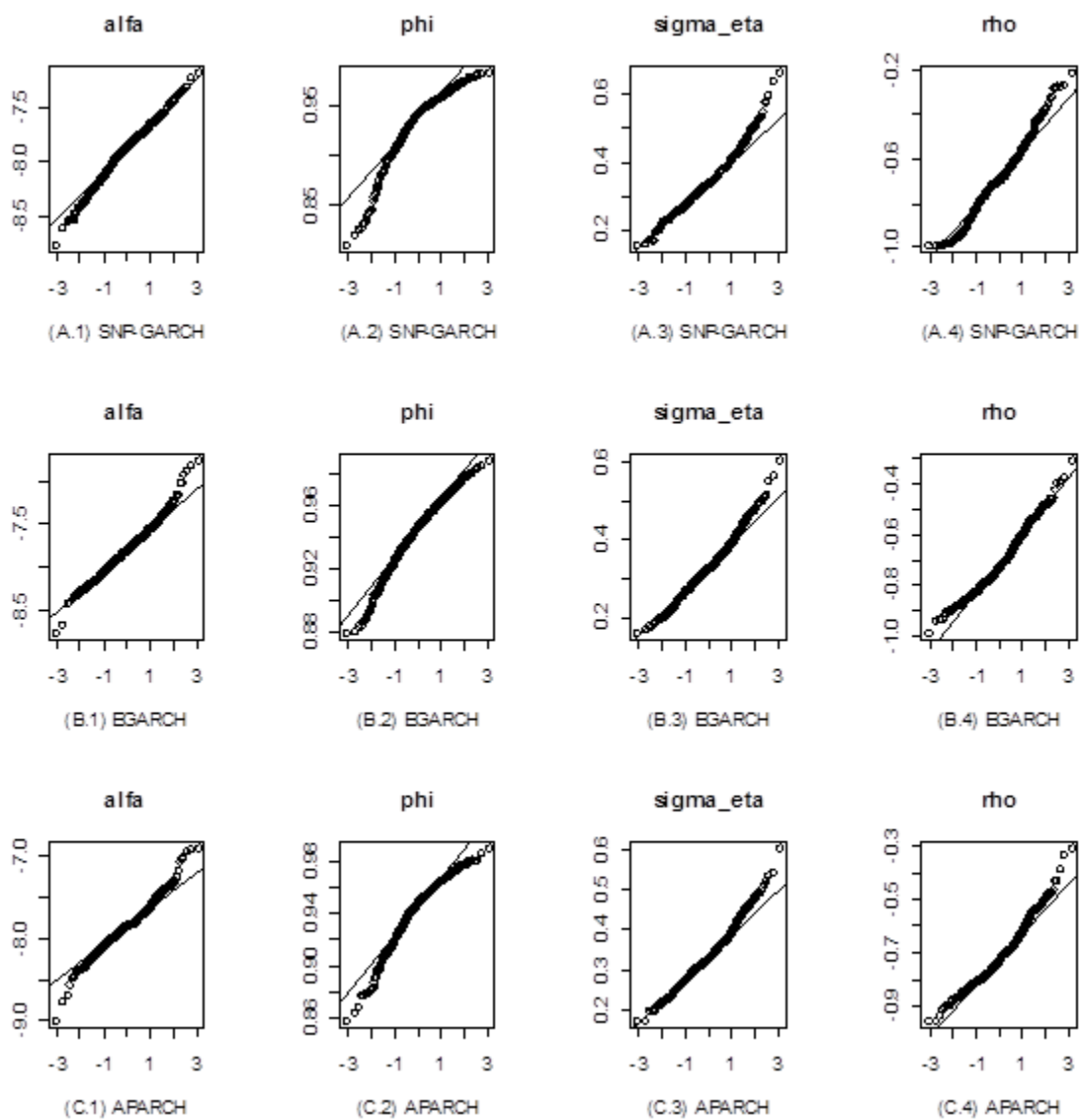


Figura 4.4: QQ-plots normales de las distribuciones de los estimadores EMM para el tamaño de muestra 500. Estos están calculados con base en 500 estimaciones convergentes dentro del espacio paramétrico del modelo de volatilidad estocástica con parámetros $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

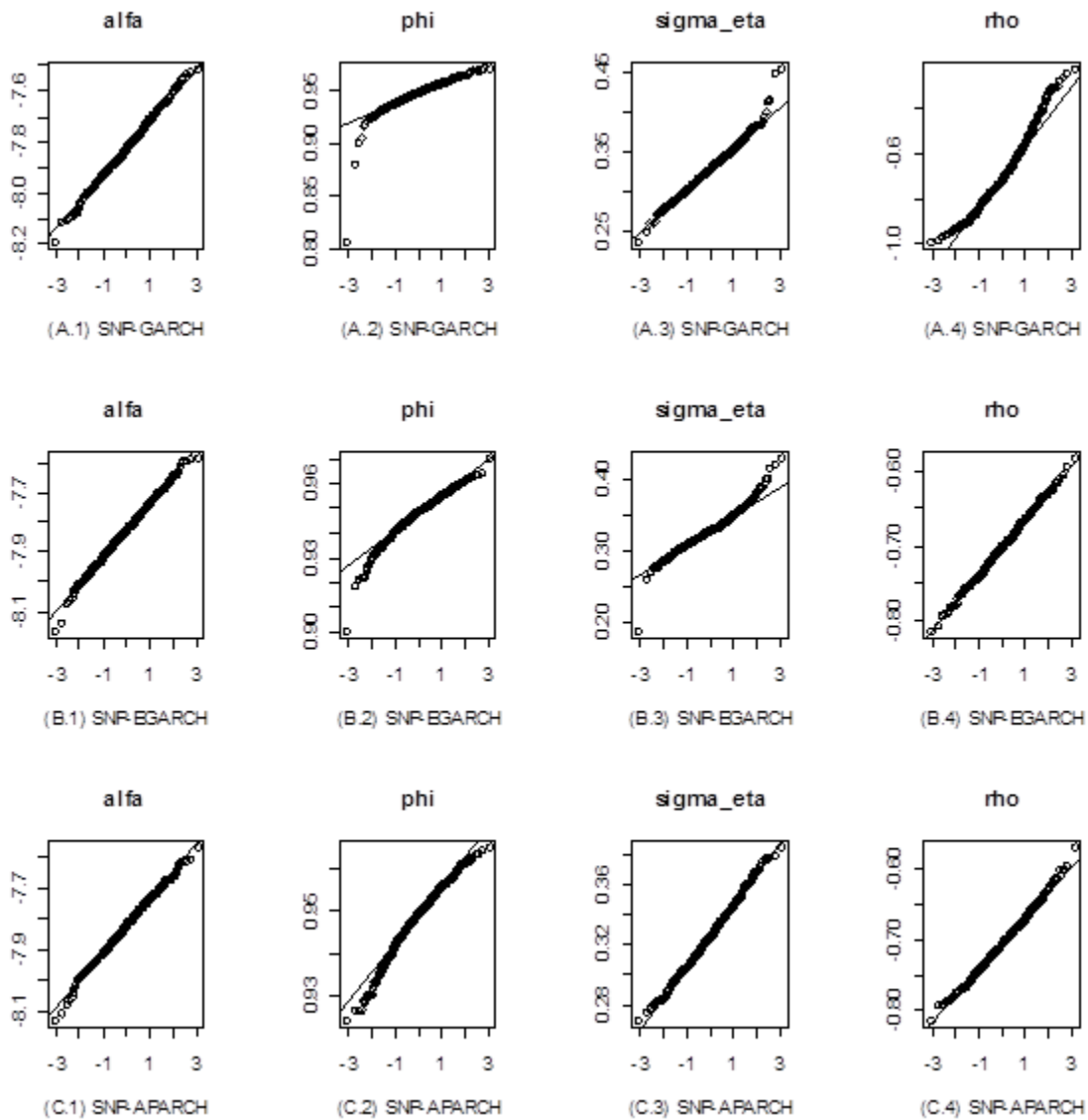


Figura 4.5: QQ-plots normales de las distribuciones de los estimadores EMM para el tamaño de muestra 4.000. Estos están calculados con base en 500 estimaciones convergentes dentro del espacio paramétrico del modelo de volatilidad estocástica con parámetros $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

Para verificar la propiedad de normalidad asintótica de los estimadores EMM, se reportan los valores P de la prueba de normalidad de Shapiro-Wilk en la tabla

4.2, y en las figuras 4.4 y 4.5 se muestran los QQ-plot normales para los estimadores con tamaño de muestra 500 y 4.000, respectivamente. Las gráficas para los tamaños de muestra 1.000 y 2.000 pueden ser consultadas en el apéndice (figuras A.3 y A.4, respectivamente). En el tamaño de muestra 500, ningún estimador ha convergido a la normalidad, según el test S-W. No obstante, en la figura 4.4 se aprecia un ajuste relativamente bueno de $\hat{\alpha}$ bajo el modelo auxiliar SNP-GARCH y bajo los modelos EGARCH y APARCH, si se excluyen las colas. Las distribuciones de los demás parámetros están lejos de la normalidad en este tamaño de muestra. La distribución del estimador EMM de α ya ha convergido a la normalidad bajo el modelo auxiliar SNP-EGARCH en muestras de tamaño 1.000 y a partir de 2.000 en los demás casos. Con respecto a la distribución de $\hat{\sigma}_\eta$, se observa una buena aproximación a la normalidad de los estimadores bajo los modelos auxiliares SNP-EGARCH y SNP-APARCH a partir del tamaño de muestra 1.000 y para el modelo auxiliar SNP-GARCH a partir de 2.000. En cuanto a $\hat{\rho}$, su distribución es normal a partir del tamaño de muestra 2.000, bajo los modelos auxiliares SNP-EGARCH y SNP-APARCH, pero para las estimaciones bajo el modelo auxiliar SNP-GARCH, no se garantiza la normalidad ni siquiera con muestras de tamaño 4.000. Por su parte, la distribución del estimador de ϕ es la que menos se aproxima a la normalidad, y solo se aprecia una buena aproximación para el tamaño de muestra 4.000. En síntesis, las inferencias basadas en la normalidad asintótica son altamente confiables con tamaños de muestra de 2.000 para los parámetros ρ y σ_η , 4.000 en el caso de ϕ y menos de 1.000 en el caso de α .

4.3.2. Modelo SV con colas pesadas

Se estimó un modelo donde la innovación de la ecuación de retornos sigue una distribución t de cola pesada. Los parámetros de este modelo son los mismos del caso anterior, pero en este modelo se tiene que

$$\eta_t = \rho\varepsilon_{t-1} + \sqrt{1 - \rho^2}\epsilon_t, \quad (4.25)$$

con $\epsilon_t \sim \text{iid}N(0, 1)$ y ε_{t-1} sigue una distribución t escalada (de media cero y varianza unitaria) con 3 grados de libertad, y ambas son mutuamente independientes. Durante el proceso de estimación EMM, no se estiman los grados de libertad sino que directamente se asumen iguales a 3. Por supuesto que en este caso, la innovación $\{\eta_t\}$ ya no es Gaussiana, sino que tiene una distribución intermedia entre la t y la normal

(Nason, 2006). Por lo tanto, $\{\sigma_t^2\}$ ya no sigue una distribución log-normal, y el vector $(\eta_t, \varepsilon_{t-1})$ no es normal bivariado, aunque la correlación entre ambas variables sigue siendo ρ . Mediante EMM es posible estimar este modelo SV, puesto que no se requiere conocer la expresión para la función de distribución de σ_t^2 .

Al igual que en la estimación del modelo SV log-normal, se llevó a cabo un estudio preliminar de modelos auxiliares, encontrándose que los que arrojan los mejores resultados son el AR(1)-EGARCH(1,1)-SNP(4) y AR(1)-APARCH(1,1)-SNP(4). En esta evaluación de los modelos auxiliares se encontró que incluir menos términos SNP conduce a tasas de falla en la convergencia más altas, mientras que modelos auxiliares con más términos SNP tienen, hasta cierto punto, menores tasas de colapso, pero las distribuciones de los estimadores se deterioran considerablemente. En este caso, la elección del modelo auxiliar es incluso más sensible para la estimación EMM que en el caso log-normal.

Por otro lado, se encontró que con tamaños de muestra menores a 2.000, las estimaciones bajo cualquier modelo auxiliar considerado arrojaron tasas de colapsos mayores al 30 %, lo cual es inconcebible, razón por la cual solo se reportaron los resultados de las estimaciones para tamaños de muestra 2.000 y 4.000. En la tabla 4.3 se presentan las principales estadísticas y en la figura 4.6 se muestran los resultados para el tamaño de muestra 4.000. Los resultados para el tamaño de muestra 2.000 se encuentran en la figura A.5 del apéndice.

Tam. muestra	2.000				4.000			
Mod. Auxiliar	AR(1)-EGARCH(1,1)-SNP(4)				AR(1)-EGARCH(1,1)-SNP(4)			
Coefficientes	α	ϕ	σ_η	ρ	α	ϕ	σ_η	ρ
Promedio	-7.886	0.932	0.326	-0.693	-7.860	0.932	0.330	-0.688
RMSE	0.183	0.040	0.063	0.089	0.140	0.037	0.053	0.082
Desv. Estándar	0.167	0.036	0.063	0.089	0.131	0.032	0.053	0.081
Valor P test S-W	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Colapsos	18.6 %				12.1 %			
Mod. Auxiliar	AR(1)-APARCH(1,1)-SNP(4)				AR(1)-APARCH(1,1)-SNP(4)			
Coefficientes	α	ϕ	σ_η	ρ	α	ϕ	σ_η	ρ
Promedio	-7.928	0.946	0.329	-0.701	-7.883	0.946	0.327	-0.706
RMSE	0.228	0.026	0.078	0.110	0.150	0.024	0.062	0.085
Desv. Estándar	0.196	0.026	0.078	0.110	0.131	0.023	0.062	0.085
Valor P test S-W	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Colapsos	15.5 %				13.8 %			

Tabla 4.3: Resultados de las estimaciones EMM del modelo de volatilidad estocástica con colas pesadas, para tamaños de muestra de 2.000 y 4.000. Los parámetros verdaderos son $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$. La cuarta fila en cada bloque corresponde al valor P del test de normalidad de Shapiro-Wilk. Los resultados están basados en 500 estimaciones convergentes.

De los resultados reportados en la tabla 4.3, el primer aspecto que salta a la vista son las tasas relativamente altas de fallas en la convergencia de los estimadores, que son superiores al 10 % incluso en el tamaño de muestra 4.000. Adicionalmente, cuando se comparan los RMSE con los del modelo log-normal de la tabla 4.2, es evidente que en el caso del modelo con colas pesadas, los RMSE son considerablemente más altos en todos los casos, y especialmente los del parámetro ϕ . En definitiva, los estimadores EMM del modelo con colas pesadas son menos precisos, además de que las tasas de falla son más altas. De la figura 4.6 se desprende que, si bien no hay grandes sesgos en los estimadores, hay una gran dispersión y una cantidad considerable de valores extremos, lo cual es evidente si se compara con el caso log-normal (figura 4.2). En cuanto al término que lidera la heteroscedasticidad condicional en el modelo auxiliar, no hay un claro dominador, pues si bien el modelo con especificación APARCH ofrece una mejor estimación de ϕ , en los demás parámetros no supera al modelo con especificación EGARCH.

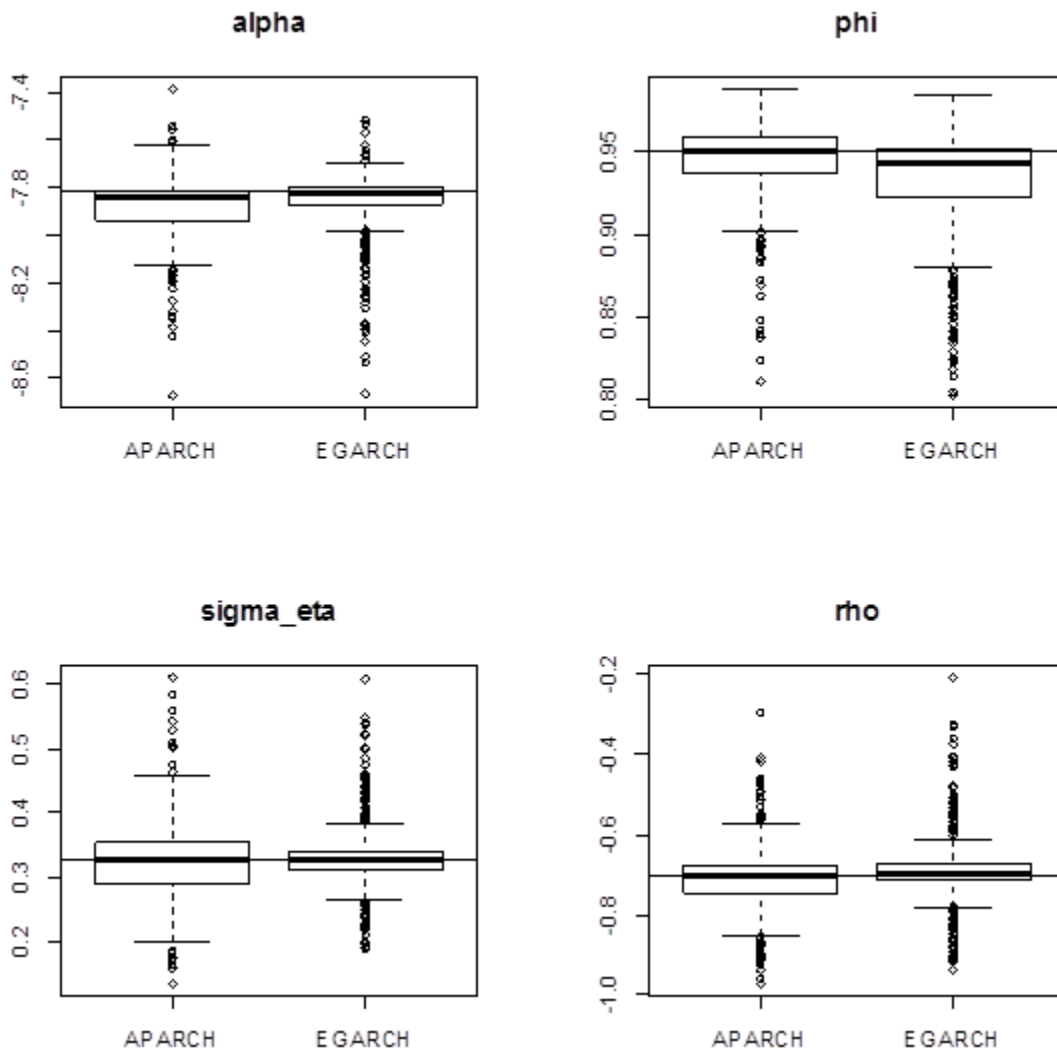


Figura 4.6: Resultados de las estimaciones EMM del modelo SV con colas pesadas, para el tamaño de muestra 4.000. Estos resultados están basados en 500 estimaciones convergentes dentro del espacio paramétrico. Las líneas horizontales corresponden a los verdaderos valores de los parámetros, que son $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

En la figura 4.7 se presentan los QQ-plot normales de los estimadores EMM para el tamaño de muestra 4.000. En la figura A.6 del apéndice se encuentra el gráfico para el tamaño de muestra 2.000. En esta figura se aprecia claramente que las distribuciones

de los estimadores están lejos de la normalidad asintótica aun con un tamaño de muestra de 4.000, el cual garantiza una muy buena aproximación en la estimación del modelo log-normal. Para el caso de colas pesadas, las distribuciones de los estimadores EMM son leptocúrticas alrededor de los verdaderos valores, pero también tienen una proporción grande de valores extremos. Estos resultados indican que no son confiables las inferencias basadas en la normalidad asintótica en las estimaciones de modelos SV con colas pesadas, bajo la estructura de estimación aquí considerada.

No obstante, hay evidencia de que las falencias en la estimación se deben a deficiencias de los modelo auxiliares utilizados para captar las principales características de este tipo de datos con colas pesadas y valores extremos de los retornos. Por un lado, cuando se estimó el modelo log-normal se observó que las estimaciones basadas en modelos SNP-GARCH, que no reproducen el efecto leverage, generaron tasas más altas de soluciones de esquina y mayor dispersión de los estimadores. Esto mismo puede estar sucediendo en este caso, es decir, que las altas tasas de colapsos se deban a que los modelos auxiliares no alcanzan a explicar las colas pesadas de los retornos. Además, el hecho de que incrementar o reducir los términos SNP en el modelo auxiliar afecte tanto los estimadores, hace sospechar que estos modelos auxiliares no son adecuados, lo que conduce al fallo de EMM. Como es enfatizado por Gallant y Long (1997), si el modelo auxiliar no es bueno, la inversa de la matriz de covarianzas asintótica (3.5) no es una matriz de ponderaciones eficiente, lo cual haría aparecer problemas de estimación similares a los que surgen con el método GMM. Para remediar estos problemas será necesario implementar una matriz de ponderaciones más elaborada o recurrir a modelos auxiliares alternativos a la densidad SNP, que cumplan las condiciones de modelar datos extremos, patrones de varianza condicional y efecto leverage y que además dispongan de una expresión analítica de la función de log-verosimilitud. Evidentemente, ambas alternativas no son una tarea sencilla.

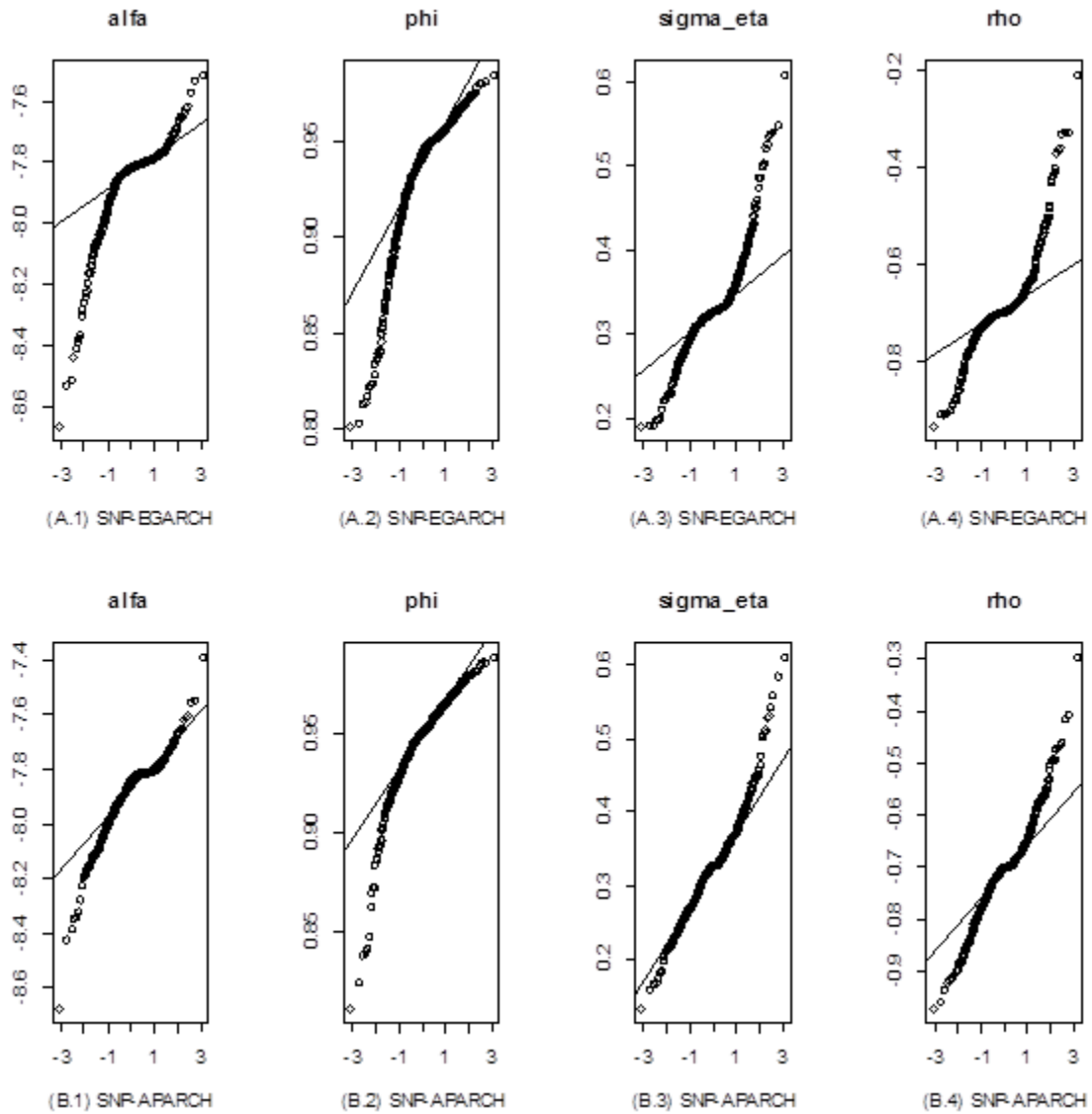


Figura 4.7: QQ-plots normales de las distribuciones de los estimadores EMM del modelo SV con colas pesadas, para el tamaño de muestra 4.000. Estos están calculados con base en 500 estimaciones convergentes dentro del espacio paramétrico del modelo de volatilidad estocástica con parámetros $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

4.4. Síntesis

En el estudio Monte Carlo realizado en este capítulo, se encontró que el EMM es un método razonablemente bueno para estimar modelos SV log-normales con efecto leverage, siempre y cuando se haga una cuidadosa selección del modelo auxiliar. A partir de un tamaño de muestra de 1.000 se obtienen estimadores precisos y con tamaños de muestra de 2.000, la distribución de los estimadores muestra una muy buena aproximación a la normalidad, por tanto las inferencias son confiables. Sin embargo, para algunos parámetros los estimadores no convergen a la normalidad asintótica a la misma velocidad. El parámetro α converge más rápidamente, incluso con solo 500 datos, mientras que ϕ lo hace más lentamente. En cuanto al parámetro de leverage, las estimaciones son precisas y la distribución asintótica se comprobó que es normal.

De forma interesante, como se presuponía con base en los resultados del estudio comparativo de modelos auxiliares de la sección 1.2, se demostró que es crucial que el modelo auxiliar contenga el efecto leverage para una mejor estimación EMM, ya que reduce las tasas de colapsos a soluciones de esquina y mejora la precisión de los estimadores. Los modelos auxiliares con términos APARCH y EGARCH son altamente recomendados, aunque los modelos con términos APARCH probaron arrojar estimadores más estables en comparación con los modelos auxiliares con términos EGARCH, a lo cual hay que agregar que esta última clase de modelos puede fallar en la estimación de la varianza condicional debido a las características de la ecuación recursiva, tal como se apreció en la sección 1.2.

En el caso de la estimación de un sencillo modelo SV con colas pesadas y efecto leverage, el desempeño del EMM no es tan satisfactorio. Si bien produce estimadores con poco sesgo, la tasa de fallas en la convergencia es alta, al igual que los RMSE, en comparación con el caso log-normal. Además, no se garantiza la convergencia de los estimadores a la distribución normal ni siquiera con tamaños de muestra del orden de 4.000, lo cual es problemático para las inferencias. Sin embargo, la evidencia sugiere que estos problemas pueden ser adjudicados a que los modelos auxiliares considerados no logran reproducir el comportamiento de los valores extremos de los retornos generados por estos procesos de volatilidad estocástica, de manera que no se logra obtener una estimación precisa de la matriz de covarianzas asintótica (3.5), afectando el algoritmo EMM.

Como fue sugerido por uno de los jurados, posteriormente sería interesante extender este estudio con un análisis de la precisión de la prueba de especificación del modelo, dada en (3.9), viendo cómo se comporta en estimaciones con base en modelos auxiliares bien y mal especificados. Esto con el objetivo de establecer si esta prueba es robusta a la elección del modelo auxiliar.

Capítulo 5

Conclusiones

Los modelos de volatilidad estocástica (SV) son versátiles para acomodar las características observadas en las series de retornos financieros, razón por la que actualmente son ampliamente utilizados en la literatura financiera. Sin embargo, su estimación no es sencilla, al no estar disponible una expresión analítica de la función de verosimilitud, lo que ha llevado a que se propongan métodos de estimación alternativos, uno de ellos es el método de momentos eficiente (EMM) de Gallant y Tauchen (1996). Para hallar los estimadores, este método utiliza los valores esperados de los scores de un modelo auxiliar bajo el modelo de volatilidad estocástica como condiciones de momentos en un criterio GMM.

En esta tesis se llevó a cabo un estudio Monte Carlo sobre la capacidad del EMM para estimar modelos de volatilidad estocástica con efecto leverage y colas pesadas en la distribución de la innovación de la ecuación de retornos. Se construyó sobre la base de los trabajos pioneros de Andersen, Chung y Sørensen (1999) y Chumacero (1997), quienes estudiaron las propiedades de los estimadores EMM en modelos SV log-normales sin efecto leverage. Aquí se amplió al caso del modelo con efecto leverage y también se puso a prueba un modelo con cola pesada en la ecuación de retornos. Ambos, el efecto leverage y las colas pesadas, han sido ampliamente documentados en la evidencia empírica de las series de retornos financieros, y de allí la importancia de determinar la capacidad del EMM para estimar modelos SV que incluyen estas propiedades.

El EMM mostró ser un método viable para la estimación del modelo SV log-normal

con leverage, produciendo estimadores insesgados y con una dispersión aceptable, siempre y cuando se haga una cuidadosa selección del modelo auxiliar. Se encontró que las estimaciones basadas en modelos auxiliares con especificaciones GARCH, que no reproducen el efecto leverage, tienen mayores tasas de falla en la convergencia, RMSE más altos y distribuciones que convergen más lentamente a la normalidad, en comparación con las estimaciones basadas en modelos auxiliares que tienen en cuenta este efecto. En este sentido, los modelos auxiliares SNP-APARCH y SNP-EGARCH son recomendados, aunque se detectó cierta inestabilidad en la ecuación recursiva del EGARCH que genera que la varianza condicional reconstruida se salga de control en algunos casos, afectando la estimación EMM. Adicionalmente, las estimaciones basadas en el modelo auxiliar SNP-APARCH mostraron ser más estables a medida que cambia el tamaño de muestra.

En la estimación EMM del modelo SV log-normal, las inferencias sobre los parámetros basadas en la normalidad asintótica son confiables con tamaños de muestra de alrededor de 2.000, aunque para el parámetro α la normalidad se alcanza incluso en tamaños de muestra menores a 1.000. Con un tamaño de muestra de 500, los estimadores son relativamente precisos e insesgados, aunque no se garantiza la convergencia a la normalidad, por lo que puede haber problemas con las inferencias. Con respecto al coeficiente de leverage, el EMM arrojó estimadores precisos y asintóticamente normales.

El desempeño del método, sin embargo, se deterioró notablemente cuando se estimó el modelo SV con colas pesadas. Aumentaron los RMSE y la dispersión de los estimadores, hubo altas tasas de falla en la convergencia, incluso con muestras grandes, y no se alcanzó la convergencia a la normalidad ni siquiera en el tamaño de muestra de 4.000. Por otro lado, los estimadores también fueron más sensibles a la elección del modelo auxiliar en comparación con el caso log-normal, ya que incluir o reducir el número de términos SNP condujo a un incremento sustancial en la tasa de colapsos o a un empeoramiento de las distribuciones de los estimadores. No obstante, estos problemas de estimación pueden ser adjudicados a que los modelos auxiliares considerados tal vez no pueden reproducir adecuadamente las características de los datos, por ende produciendo malos estimadores de la matriz de ponderaciones. En estas circunstancias, se requerirá acudir a matrices de ponderaciones más elaboradas o buscar modelos auxiliares alternativos a SNP, los cuales puedan explicar las colas pesadas de los retornos, a la vez que incluyan el efecto leverage, el comportamiento

persistente de la varianza condicional y que permitan obtener expresiones analíticas para los scores. Claramente, encontrar modelos auxiliares con estas características no es una tarea fácil.

La principal ventaja del EMM, que es recurrir a un modelo auxiliar, se convierte en un inconveniente cuando no es posible hallar un modelo auxiliar que describa bien los datos, como ocurre cuando los retornos son generados por modelos SV con colas pesadas o cuando la varianza condicional es menos persistente, en cuyo caso los modelos de heteroscedasticidad condicional no ofrecen una buena aproximación, tal como se vio en la sección 1.2. Recurrir a modelos auxiliares inadecuados hace que el EMM sea más impreciso, pudiendo errar considerablemente en los parámetros estimados o conducir a rechazar o aceptar un modelo SV incorrectamente. En estas situaciones, el investigador queda no solo con el problema de determinar cuál modelo SV ajustar a los datos, sino que también debe asumir el reto de hallar un modelo auxiliar adecuado para la estimación.

En lo que respecta a la aplicabilidad del método EMM en situaciones reales, se puede mencionar lo siguiente. Por un lado, los resultados de las estimaciones permiten deducir que la precisión del método es cuestionable en situaciones en las que los datos proceden de modelos no Gaussianos o con especificaciones más complejas que las consideradas en este estudio. Pero, se debe recalcar, el método no falla en su concepción teórica, sino en la dificultad práctica de hallar un modelo auxiliar que filtre adecuadamente las características de los datos. Esta alta dependencia del modelo auxiliar hace al EMM un método riesgoso en manos de usuarios no expertos, ya que una mala elección de éste puede conducir a errores graves de estimación y a inferencias erróneas sobre los modelos estimados. Esto es bien ilustrado por Zivot y Wang (2006, pp. 957-959), quienes estiman un modelo SV básico recurriendo a un inadecuado modelo auxiliar AR(1) y a un modelo auxiliar más elaborado, que explica mejor las características de los datos. En la primera estimación, el criterio chi-cuadrado del EMM sugiere que el modelo básico se ajusta bien a los datos, pero, al hacer la estimación con el modelo auxiliar adecuado, esta hipótesis es rechazada contundentemente. Por otro lado, el asunto del modelo auxiliar también puede ser un factor clave que determine la aceptación del método por parte de los usuarios. Si bien una vez implementado en software, el EMM es fácil de aplicar, es menester del usuario elegir un modelo auxiliar para representar los datos. Este aspecto, seguramente, haría que muchos usuarios no especializados opten por otros métodos de estimación o que

finalmente recurran a modelos de tipo GARCH, más fácilmente estimables.

Bibliografía

- [1] Andersen, T. G., y Sørensen, B. E. (1996), “GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study,” *Journal of Business & Economic Statistics*, 14, 328-352.
- [2] Andersen, T. G., Chung, H.-J., y Sørensen, B. E. (1999), “Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study,” *Journal of Econometrics*, 91, 61-87.
- [3] Andersen, T. G., Bollerslev, T., Diebold, F. X., y Ebens, H. (2001), “The Distribution of Realized Stock Return Volatility,” *Journal of Financial Economics*, 61, 43-76.
- [4] Andersen, T. G., Bollerslev, T., Diebold, F. X., y Labys, P. (2001), “The Distribution of Realized Exchange Rate Volatility,” *Journal of the American Statistical Association*, 96, 42-55.
- [5] Andersson, J. (2001), “On the Normal Inverse Gaussian Stochastic Volatility Model,” *Journal of Business & Economic Statistics*, 19, 44-54.
- [6] Bansal, R., Gallant, A. R., Hussey, R., y Tauchen, G. E. (1995), “Nonparametric Estimation of Structural Models for High-Frequency Currency Market Data,” *Journal of Econometrics*, 66, 251-287.
- [7] Barndorff-Nielsen, O. E. (1997), “Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling,” *Scandinavian Journal of Statistics*, 24, 1-13.
- [8] Barndorff-Nielsen, O. E., y Shephard, N. (2002), “Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 253-280.

- [9] Black, F. (1976), “Studies in Stock Price Volatility Changes,” en: *American Statistical Association Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section*, pp. 177-181.
- [10] Black, F., y Scholes, M. (1973), “The Pricing of Options and Corporate Liabilities,” *Journal of Political Economy*, 81, 637-654.
- [11] Bollerslev, T. (1986), “Generalized Autoregressive Conditional Heteroskedasticity,” *Journal of Econometrics*, 31, 307-327.
- [12] Bollerslev, T., Litvinova, J., y Tauchen, G. (2006), “Leverage and Volatility Feedback Effects in High-Frequency Data,” *Journal of Financial Econometrics*, 4, 353-384.
- [13] Brockwell, P. J., y Davis, R. A. (1991), *Time Series: Theory and Methods*, segunda edición, Nueva York: Springer-Verlag.
- [14] Broto, C., y Ruiz, E. (2004), “Estimation Methods for Stochastic Volatility Models: A Survey,” *Journal of Economic Surveys*, 18, 613-649.
- [15] Carnero, M. A., Peña, D., y Ruiz, E. (2004), “Persistence and Kurtosis in GARCH and Stochastic Volatility Models,” *Journal of Financial Econometrics*, 2, 319-342.
- [16] Chernov, M., Gallant, R., Ghysels, E., y Tauchen, G. (2003), “Alternative Models for Stock Price Dynamics,” *Journal of Econometrics*, 116, 225-257.
- [17] Chib, S., Nardari, F., y Shephard, N. (2002), “Markov Chain Monte Carlo Methods for Stochastic Volatility Models,” *Journal of Econometrics*, 108, 281-316.
- [18] Chumacero, R. A. (1997), “Finite Sample Properties of the Efficient Method of Moments,” *Studies in Nonlinear Dynamics & Econometrics*, 2, 35-51.
- [19] Clark, P. K. (1973), “A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices,” *Econometrica*, 41, 135-155.
- [20] Cont, R. (2001), “Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues,” *Quantitative Finance*, 1, 223-236.

- [21] Davis, R. A., y Mikosch, T. (2009a), “Extremes of Stochastic Volatility Models,” en: Andersen, T. G., Davis, R. A., Kreiß, J.-P., y Mikosch, T. (Eds.), *Handbook of Financial Time Series*, Berlín: Springer-Verlag, pp. 355-364.
- [22] Davis, R. A., y Mikosch, T. (2009b), “Probabilistic Properties of Stochastic Volatility Models,” en: Andersen, T. G., Davis, R. A., Kreiß, J.-P., y Mikosch, T. (Eds.), *Handbook of Financial Time Series*, Berlín: Springer-Verlag, pp. 255-267.
- [23] Ding, Z., Granger, C. W. J., y Engle, R. F. (1993), “A Long Memory Property of Stock Market Returns and a New Model,” *Journal of Empirical Finance*, 1, 83-106.
- [24] Duffie, D., y Singleton, K. J. (1993), “Simulated Moments Estimation of Markov Models of Asset Prices,” *Econometrica*, 61, 929-952.
- [25] Engle, R. F. (1982), “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation,” *Econometrica*, 50, 987-1007.
- [26] Engle, R. F., y Ng, V. K. (1993), “Measuring and Testing the Impact of News on Volatility,” *Journal of Finance*, 48, 1749-1778.
- [27] Fama, E. F. (1965), “The Behavior of Stock-Market Prices,” *Journal of Business*, 38, 34-105.
- [28] Gallant, A. R., y Long, J. R. (1997), “Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Squared,” *Biometrika*, 84, 125-141.
- [29] Gallant, A. R., y Nychka, D. W. (1987), “Seminonparametric Maximum Likelihood Estimation,” *Econometrica*, 55, 363-390.
- [30] Gallant, A. R., y Tauchen, G. (1989), “Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications,” *Econometrica*, 57, 1091-1120.
- [31] Gallant, A. R., y Tauchen, G. E. (1996), “Which Moments to Match?” *Econometric Theory*, 12, 657-681.
- [32] Gallant, A. R., y Tauchen, G. (1999), “The Relative Efficiency of Method of Moments Estimators,” *Journal of Econometrics*, 92, 149-172.

- [33] Gallant, A. R., y Tauchen, G. (2002), “Efficient Method of Moments,” *Duke Economics Working Paper [en línea]*, No. 02-06. Disponible en <http://econ.duke.edu/research/working-paper-series>
- [34] Gallant, A. R., y Tauchen, G. (2010a), “EMM: A Program for Efficient Method of Moments Estimation,” guía de usuario, versión 2.6, disponible en <http://econ.duke.edu/webfiles/arg/emm>
- [35] Gallant, A., y Tauchen, G. (2010b), “SNP: A Program for Nonparametric Time Series Analysis,” guía de usuario, versión 9.0. Disponible en: <http://www.econ.duke.edu/webfiles/arg/snp>
- [36] Gallant, A. R., Hsieh, D., y Tauchen, G. (1997), “Estimation of Stochastic Volatility Models with Diagnostics,” *Journal of Econometrics*, 81, 159-192.
- [37] Ghysels, E., Harvey, A. C., y Renault, E. (1996), “Stochastic Volatility,” en: Maddala, G. S., y Rao, C. R. (Eds.), *Handbook of Statistics 14: Statistical Methods in Finance*, Amsterdam: Elsevier, pp. 119-192.
- [38] Giraldo, N. D., y Lopera, M. (2006), “Estimación Eficiente en un Modelo de Volatilidad Estocástica con Asimetría,” documento no publicado, Universidad Nacional de Colombia-Medellín.
- [39] Glosten, L., Jagannathan, R., y Runkle, D. (1993), “On the Relation between Expected Value and the Volatility of the Nominal Excess Return on Stocks,” *Journal of Finance*, 48, 1779-1801.
- [40] Grimmett, G. R., y Stirzaker, D. R. (2001), *Probability and Random Processes*, tercera edición, Nueva York: Oxford University Press.
- [41] Hansen, L. P. (1982), “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica*, 50, 1029-1054.
- [42] Harvey, A. C. (1998), “Long Memory in Stochastic Volatility,” en: Knight, J. L., y Satchell, S. (Eds.), *Forecasting Volatility in the Financial Markets*, Oxford: Butterworth-Heinemann.
- [43] Harvey, A. C., y Shephard, N. G. (1996), “The Estimation of an Asymmetric Stochastic Variance Model for Asset Returns,” *Journal of Business & Economic Statistics*, 14, 429-434.

- [44] Harvey, A. C., Ruiz, E., y Shephard, N. (1994), "Multivariate Stochastic Variance Models," *Review of Economic Studies*, 61, 247-264.
- [45] He, C., Teräsvirta, T., y Malmsten, H. (2002), "Moment Structure of a Family of First-Order Exponential GARCH Models," *Econometric Theory*, 18, 868-885.
- [46] Hull, J., y White, A. (1987), "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, 42, 281-300.
- [47] Instituto SAS (en línea), "Efficient Method of Moments Estimation of a Stochastic Volatility Model," ejemplo en sitio web. URL: <http://www.support.sas.com/rnd/app/examples/ets/emmweb/index.htm> (consultado en marzo de 2011).
- [48] Jacquier, E., Polson, N. G., y Rossi, P. E. (1994), "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business & Economic Statistics*, 12, 371-389.
- [49] Jacquier, E., Polson, N. G., y Rossi, P. E. (2004), "Bayesian Analysis of Stochastic Volatility Models with Fat-Tails and Correlated Errors," *Journal of Econometrics*, 122, 185-212.
- [50] Jondeau, E., y Rockinger, M. (2003), "Testing for Differences in the Tails of Stock-Market Returns," *Journal of Empirical Finance*, 10, 559-581.
- [51] Jungbacker, B., y Koopman, S. J. (2009), "Parameter Estimation and Practical Aspects of Modeling Stochastic Volatility," en: Andersen, T. G., Davis, R. A., Kreiß, J.-P., y Mikosch, T. (Eds.), *Handbook of Financial Time Series*, Berlín: Springer-Verlag, pp. 313-344.
- [52] Kim, S., Shephard, N., y Chib, S. (1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65, 361-393.
- [53] Liesenfeld, R., y Jung, R. C. (2000), "Stochastic Volatility Models: Conditional Normality versus Heavy-Tailed Distributions," *Journal of Applied Econometrics*, 15, 137-160.
- [54] Liesenfeld, R., y Richard, J.-F. (2003), "Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics," *Journal of Empirical Finance*, 10, 505-531.

- [55] Ling, S., y McAleer, M. (2002), “Necessary and Sufficient Moment Conditions for the GARCH(r,s) and Asymmetric Power GARCH(r,s) Models,” *Econometric Theory*, 18, 722-729.
- [56] Mandelbrot, B. (1963), “The Variation of Certain Speculative Prices,” *Journal of Business*, 36, 394-419.
- [57] Melino, A., y Turnbull, S. M. (1990), “Pricing Foreign Currency Options with Stochastic Volatility,” *Journal of Econometrics*, 45, 239-265.
- [58] Nason, G. P. (2006), “On the Sum of t and Gaussian Random Variables,” *Statistics & Probability Letters*, 76, 1280-1286.
- [59] Nelson, D. B. (1991), “Conditional Heteroskedasticity in Asset Returns: A New Approach,” *Econometrica*, 59, 347-370.
- [60] Osborne, M. F. (1959), “Brownian Motion in the Stock Market,” *Operations Research*, 7, 145-173.
- [61] Renault, E. (2009), “Moment-Based Estimation of Stochastic Volatility Models,” en: Andersen, T. G., Davis, R. A., Kreiß, J.-P., y Mikosch, T. (Eds.), *Handbook of Financial Time Series*, Berlín: Springer-Verlag, pp. 269-311.
- [62] Shephard, N., y Andersen, T. G. (2009), “Stochastic Volatility: Origins and Overview,” en: Andersen, T. G., Davis, R. A., Kreiß, J.-P., y Mikosch, T. (Eds.), *Handbook of Financial Time Series*, Berlín: Springer-Verlag, pp. 233-254.
- [63] Straumann, D., Y Mikosch, T. (2006), “Quasi-Maximum-Likelihood Estimation in Heteroscedastic Time Series: A Stochastic Recurrence Equations Approach,” *Annals of Statistics*, 34, 2449-2495.
- [64] Taylor, S. J. (1982), “Financial Returns Modelled by the Product of Two Stochastic Processes: A Study of Daily Sugar Prices, 1961-79,” en: Anderson, O.D. (Ed.), *Time Series Analysis: Theory and Practice 1*, Amsterdam: North-Holland, pp. 203-226. Reproducido en: Shephard, N. (2005), *Stochastic Volatility: Selected Readings*, Nueva York: Oxford University Press, pp. 60-82.
- [65] Taylor, S. J. (1986), *Modeling Financial Time Series*, Chichester: John Wiley and Sons.

- [66] Taylor, S. J. (1994), “Modeling Stochastic Volatility: A Review and Comparative Study,” *Mathematical Finance*, 4, 183-204.
- [67] van der Sluis, P. J. (1997), “EmmPack 1.01: C/C++ Code for Use with Ox for Estimation of Univariate Stochastic Volatility Models with the Efficient Method of Moments,” *Studies in Nonlinear Dynamics & Econometrics*, 2, 77-94.
- [68] Watanabe, T. (1999), “A non-Linear Filtering Approach to Stochastic Volatility Models with an Application to Daily Stock Returns,” *Journal of Applied Econometrics*, 14, 101-121.
- [69] Wiggins, J. B. (1987), “Option Values Under Stochastic Volatility: Theory and Empirical Estimates,” *Journal of Financial Economics*, 19, 351-372.
- [70] Yu, J. (2005), “On Leverage in a Stochastic Volatility Model,” *Journal of Econometrics*, 127, 165-178.
- [71] Zivot, E., y Wang, J. (2006), *Modeling Financial Time Series with S-Plus*, Nueva York: Springer-Verlag.

Apéndice A

Resultados de las estimaciones para tamaños de muestra 1.000 y 2.000

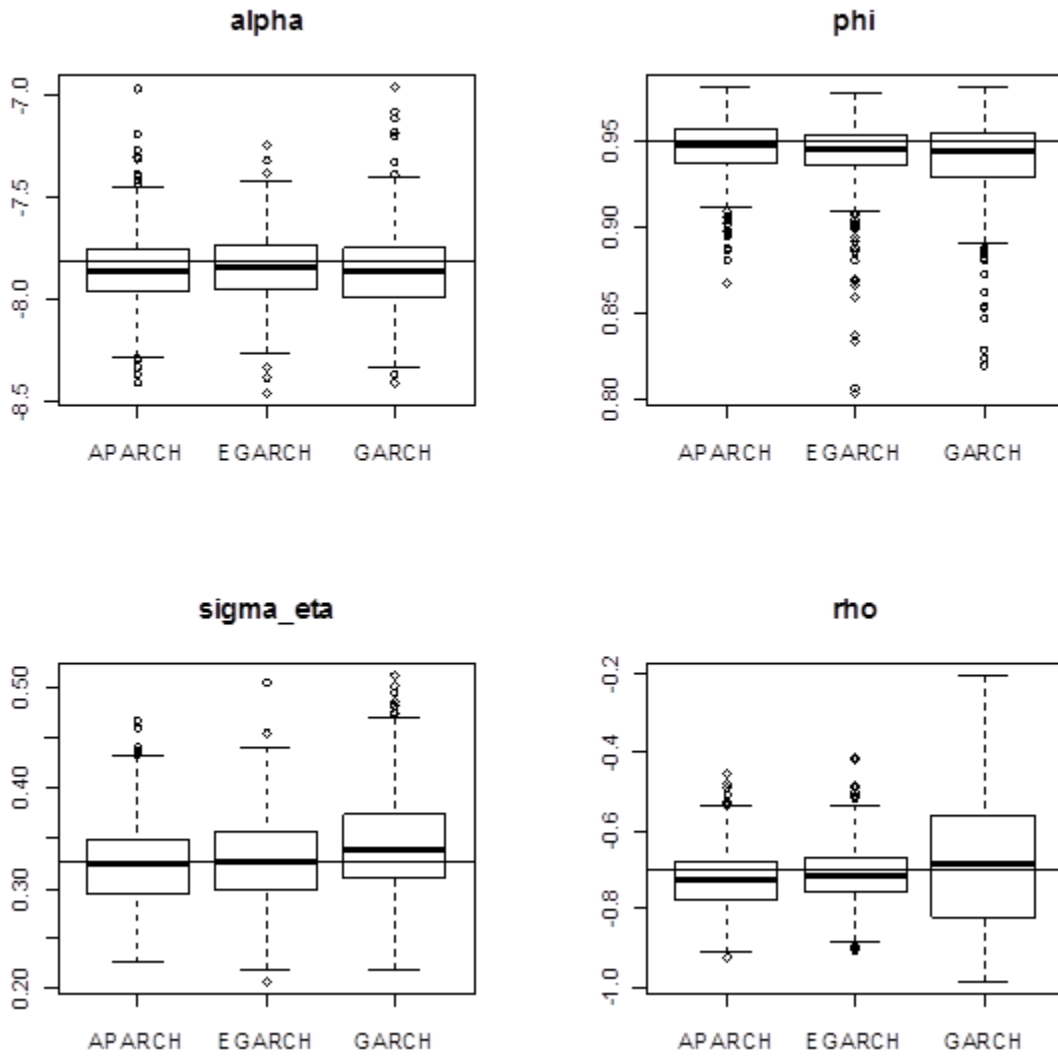


Figura A.1: Resultados de las estimaciones EMM para el tamaño de muestra 1.000. Estos resultados están basados en 500 estimaciones convergentes dentro del espacio paramétrico. Las líneas horizontales corresponden a los verdaderos valores de los parámetros, que son $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

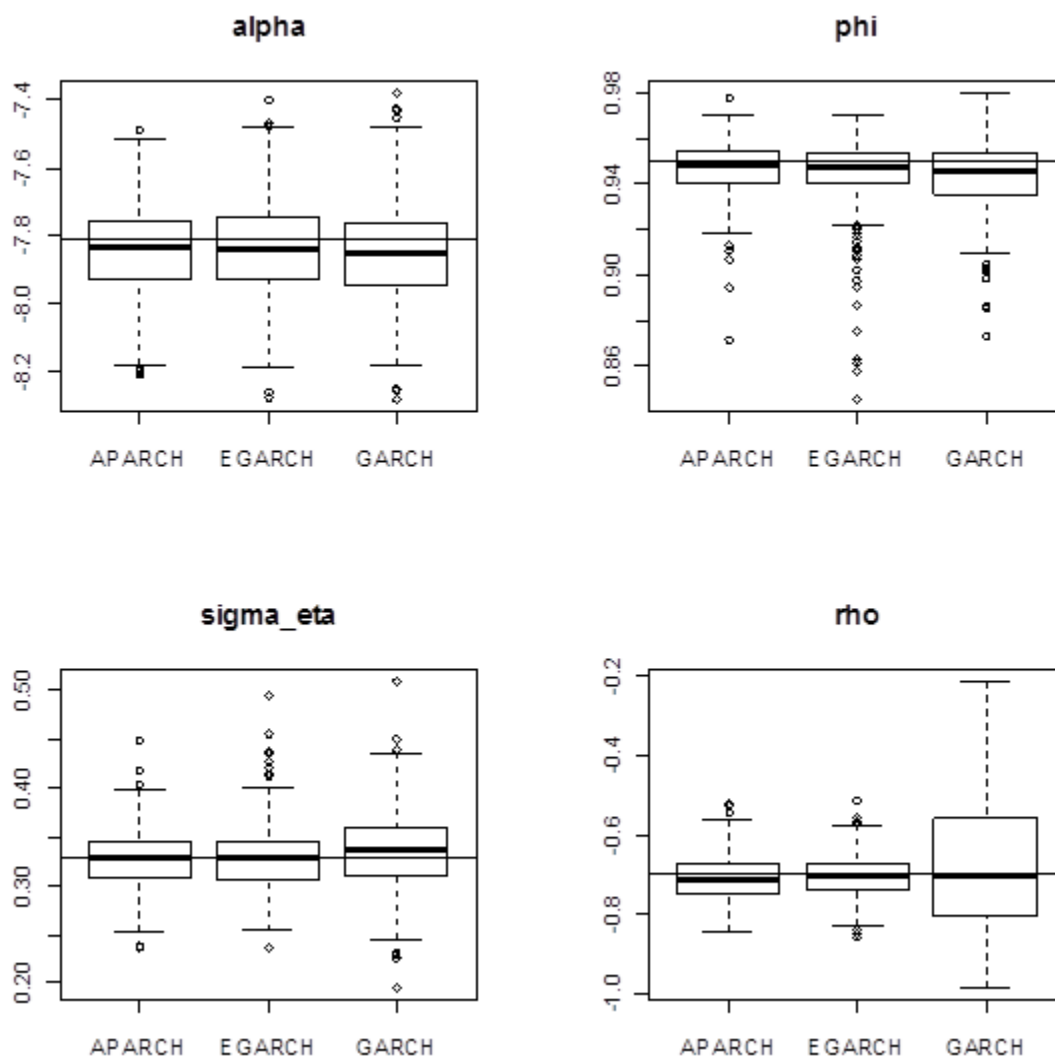


Figura A.2: Resultados de las estimaciones EMM para el tamaño de muestra 2.000. Estos resultados están basados en 500 estimaciones convergentes dentro del espacio paramétrico. Las líneas horizontales corresponden a los verdaderos valores de los parámetros, que son $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

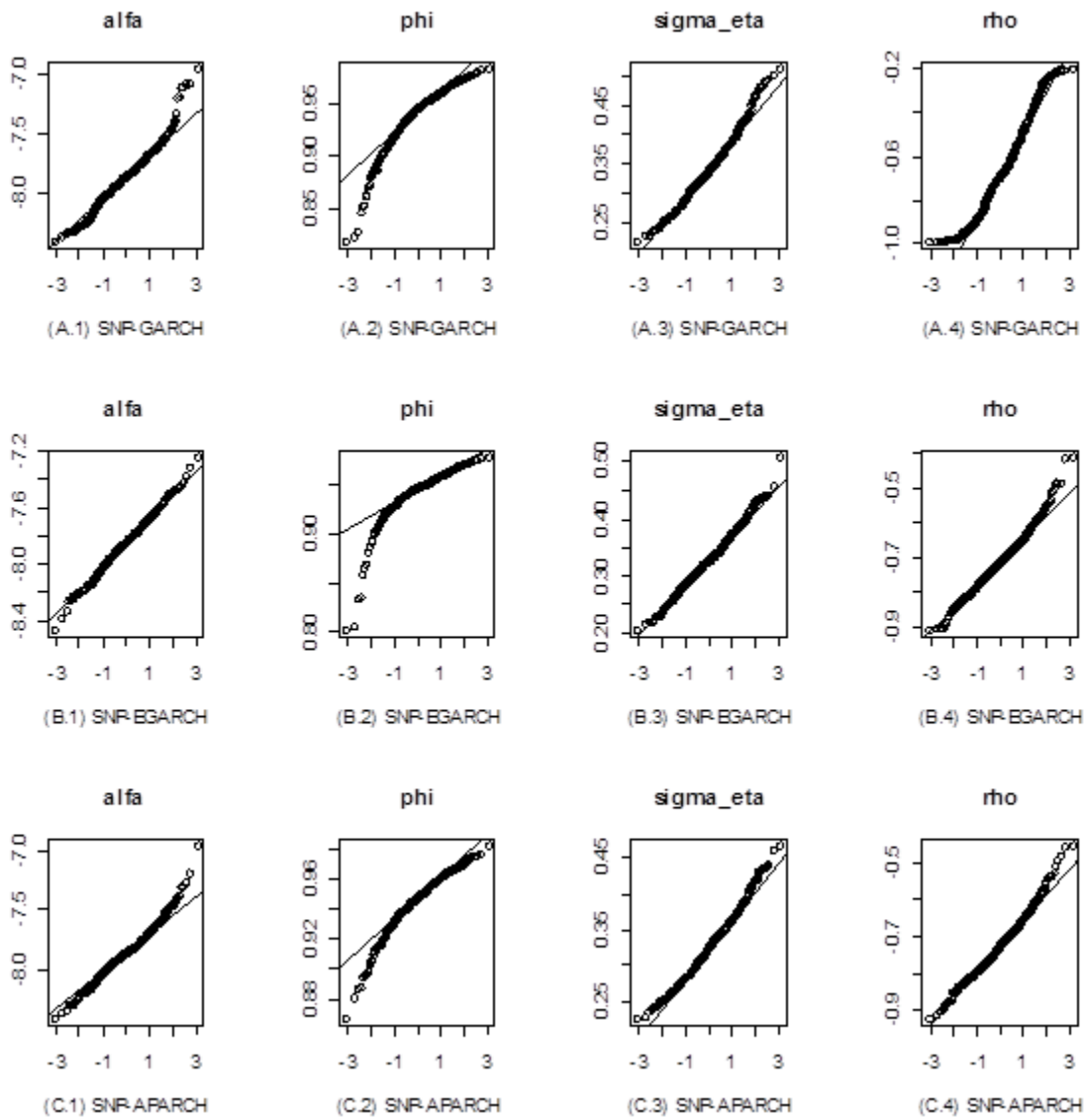


Figura A.3: QQ-plots normales de las distribuciones de los estimadores EMM para el tamaño de muestra 1.000. Estos están calculados con base en 500 estimaciones convergentes dentro del espacio paramétrico del modelo de volatilidad estocástica con parámetros $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

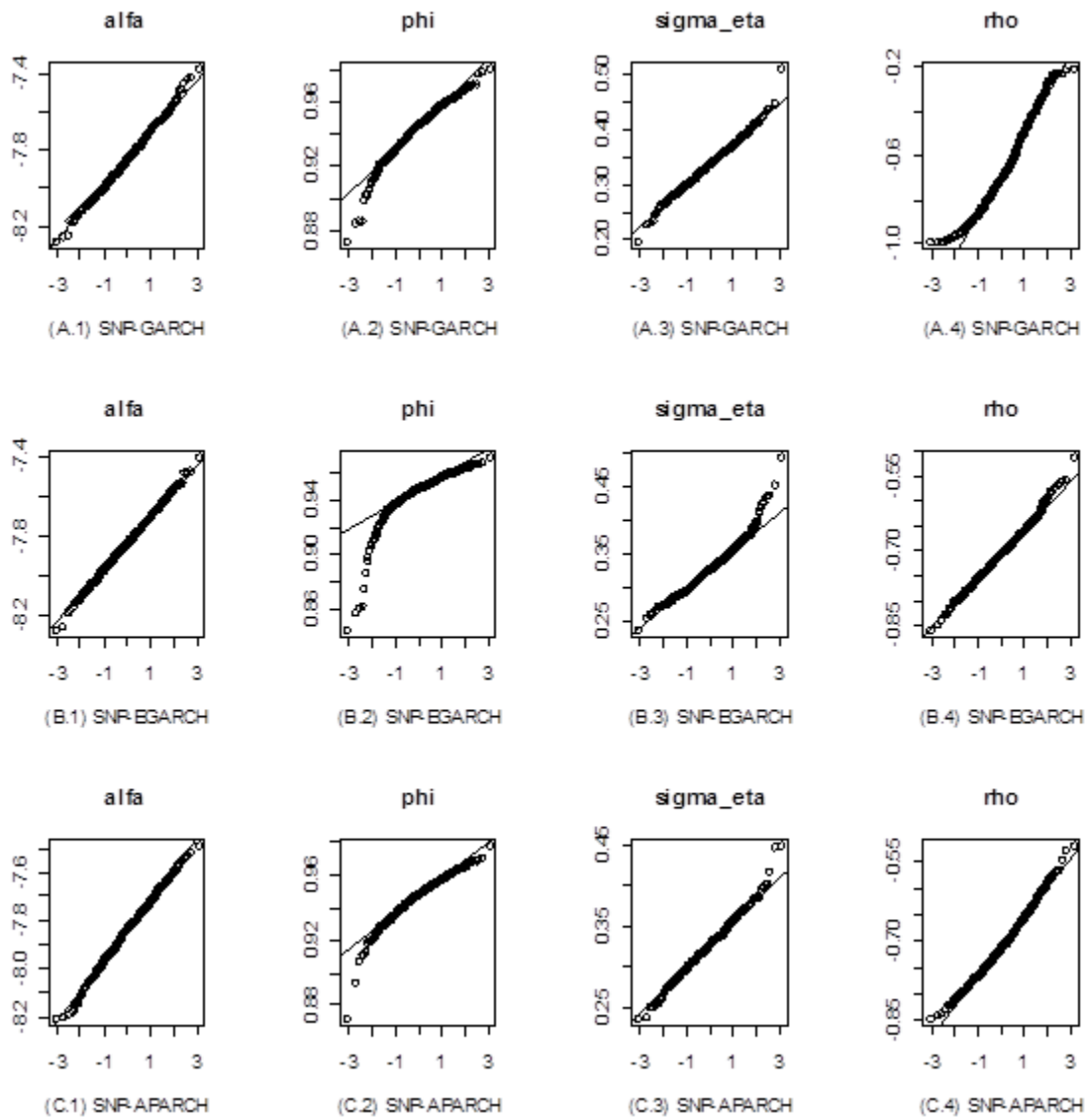


Figura A.4: QQ-plots normales de las distribuciones de los estimadores EMM para el tamaño de muestra 2.000. Estos están calculados con base en 500 estimaciones convergentes dentro del espacio paramétrico del modelo de volatilidad estocástica con parámetros $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

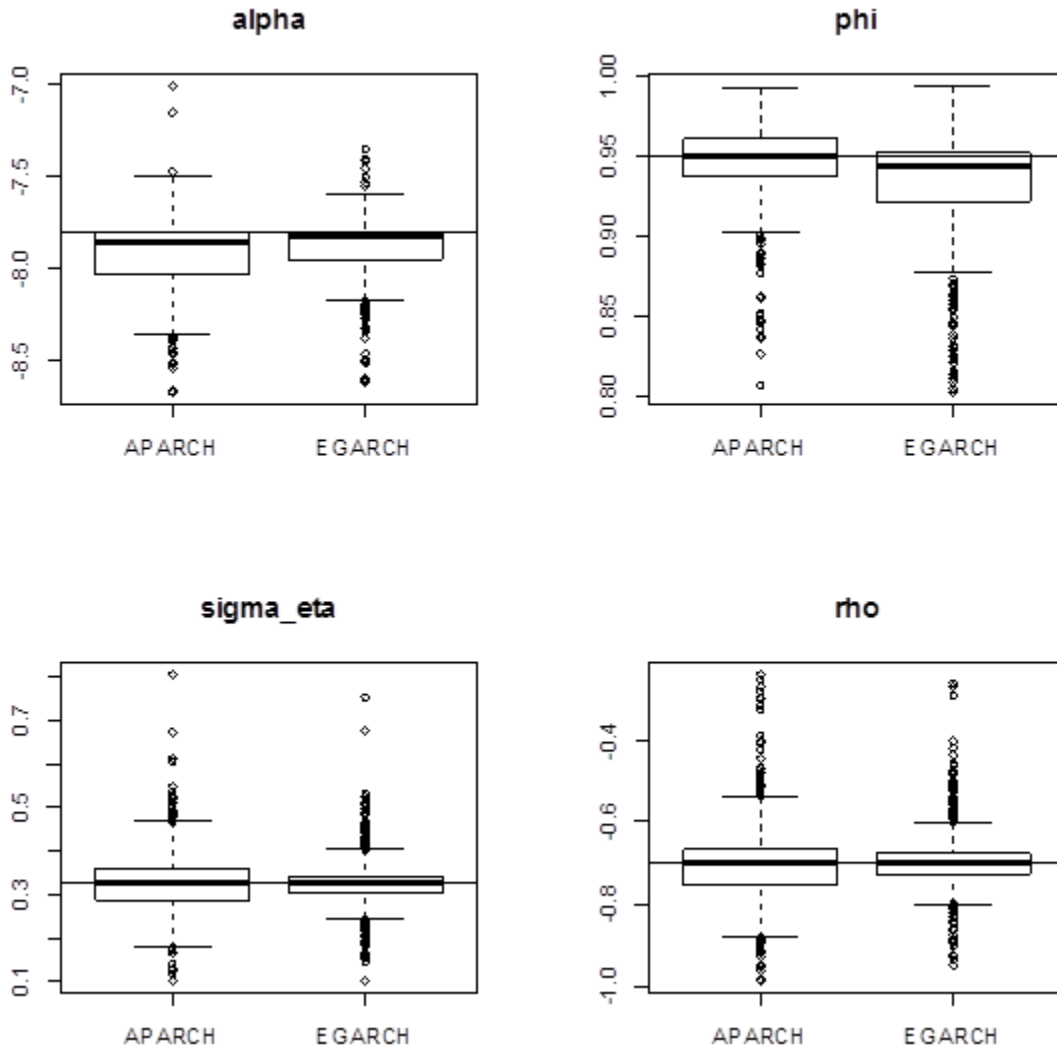


Figura A.5: Resultados de las estimaciones EMM del modelo SV con colas pesadas, para el tamaño de muestra 2.000. Estos resultados están basados en 500 estimaciones convergentes dentro del espacio paramétrico. Las líneas horizontales corresponden a los verdaderos valores de los parámetros, que son $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

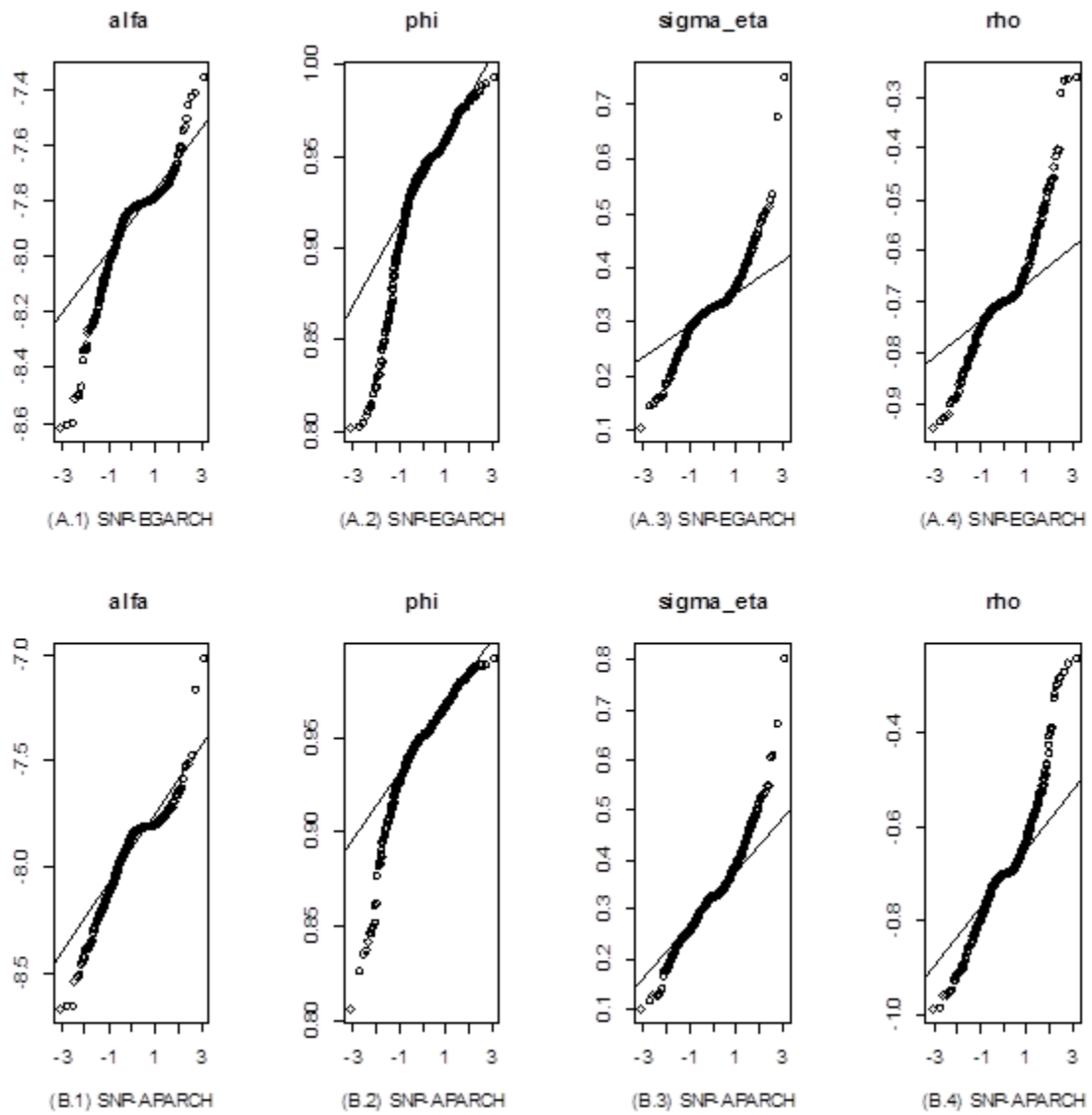


Figura A.6: QQ-plots normales de las distribuciones de los estimadores EMM del modelo SV con colas pesadas, para el tamaño de muestra 2.000. Estos están calculados con base en 500 estimaciones convergentes dentro del espacio paramétrico del modelo de volatilidad estocástica con parámetros $\alpha = -7.81$, $\phi = 0.95$, $\sigma_\eta = 0.3273$ y $\rho = -0.7$.

Apéndice B

Código en R utilizado para las estimaciones

A continuación se provee el programa en R utilizado para la estimación de las densidades SNP con términos GARCH, EGARCH y APARCH y para la estimación del modelo SV por EMM. El programa se ejecutó en la versión 2.13.2 del Software. Primero se exponen las funciones que se emplean en la estimación y al final se presenta un ejemplo. Los tiempos promedio de computación de las estimaciones EMM en un equipo con procesador Intel Core 2 Duo @2.8GHz estuvieron entre 1 minuto para la estimación con modelo auxiliar SNP-GARCH y cerca de 3 minutos en las estimaciones con modelo auxiliar SNP-APARCH en el caso del modelo SV log-normal. Para el modelo SV con colas pesadas, los tiempos fueron de cerca de 2 minutos para el modelo auxiliar SNP-EGARCH y 5 minutos para el auxiliar SNP-APARCH.

En primer lugar, se creó una librería DLL en lenguaje C que contiene las ecuaciones recursivas de los modelos GARCH(1,1), EGARCH(1,1), APARCH(1,1), así como las ecuaciones recursivas correspondientes a los scores de cada uno de los modelos. Estas funciones están guardadas en el archivo “funcionesC.dll”, y puede ser proveído a solicitud. Estas se cargan escribiendo:

```
dyn.load("funcionesC.dll")
```

Mediante la siguiente función, se generan las volatilidades. Esta es una función auxiliar, que se empleará posteriormente en la estimación SNP.

```

eps.sig.gen = function(theta,ar,y,x,arch.model){
n <- length(y)
a0 <- theta[1]
if (ar==0) mu <- a0
else {
a1 <- theta[2]
mu = a0 + a1*c((mean(x)-a0)/a1,x[-n])
}
epsilon <- y - mu
epsil.x <- x - mu

if (arch.model=="garch"){
if (ar==0){c0<-theta[2]; c1<-theta[3]; d1<-theta[4]}
else {      c0<-theta[3]; c1<-theta[4]; d1<-theta[5]}
vrz = var(x)
sigma = .C("sigmaGARCH",as.double(epsil.x),as.double(c0),
as.double(c1),as.double(d1),as.integer(n),as.double(vrz),
sigma=double(n))$sigma
}

if (arch.model=="egarch"){
if (ar==0){c0<-theta[2]; c1<-theta[3]; d1<-theta[4]; m1<-theta[5]}
else {      c0<-theta[3]; c1<-theta[4]; d1<-theta[5]; m1<-theta[6]}
logv = log(var(x))
sigma = .C("sigmaEGARCH",as.double(epsil.x),as.double(abs(epsil.x)),
as.double(c0),as.double(c1),as.double(d1),as.double(m1),
as.integer(n),as.double(logv),sigma=double(n))$sigma
}

if (arch.model=="aparch"){
if (ar==0){c0<-theta[2]; c1<-theta[3]; d1<-theta[4]; m1<-theta[5];
delta<-theta[6]}
else {      c0<-theta[3]; c1<-theta[4]; d1<-theta[5]; m1<-theta[6];
delta<-theta[7]}
}

```



```

sdX = sd(x)
sigma = .C("sigmaAPARCH",as.double(epsil.x),as.double(abs(epsil.x)),
as.double(c0),as.double(c1),as.double(d1),as.double(m1),
as.double(delta),as.integer(n),as.double(sdx),sigma=double(n))$sigma
}

sal = cbind(epsilon,sigma)
colnames(sal) = c("epsilon","sigma")
return(sal)
}

```

Utilizando la siguiente función, se generan los scores. Esta, al igual que la anterior, es una función auxiliar que se utiliza luego en la estimación de las densidades SNP.

```

SNP.score.gen = function(solution,SNPvec,bvec,z,arch.model,epsilon,
sigma,y,ar){
b = rep(NA,length(SNPvec))
for (i in 1:length(b)){
if (SNPvec[i]==0) b[i]=0
if (SNPvec[i]==1){b[i]=bvec[1]; bvec=bvec[-1]}
}
if (length(b)<8) b[(length(b)+1):8]=0

n <- length(y)
e0 = 0.001

denom = 1 + b[1]^2 + 6*b[1]*(b[3]+5*b[5]+35*b[7]) + 3*b[2]^2 +
2*b[2]*(15*b[4]+105*b[6]+945*b[8]+1) + 15*b[3]^2 + 210*b[3]*
(b[5]+9*b[7]) + 105*b[4]^2 + 6*b[4]*(315*b[6]+3465*b[8]+1) +
945*b[5]^2 + 20790*b[5]*b[7] + 10395*b[6]^2 + 30*b[6]*(9009*b[8]+
1) + 135135*b[7]^2 + 2027025*b[8]^2 + 210*b[8] + e0

poli = 1 + b[1]*z + b[2]*z^2 + b[3]*z^3 + b[4]*z^4 + b[5]*z^5 +
b[6]*z^6 + b[7]*z^7 + b[8]*z^8

```

```

dDenom.db1 = 2*b[1] + 6*(b[3]+5*b[5]+35*b[7])
if (b[2]!=0) dDenom.db2 = 6*b[2] + 2*(15*b[4]+105*b[6]+945*b[8]+1)
else dDenom.db2 = 0
dDenom.db3 = 6*b[1] + 30*b[3] + 210*(b[5]+9*b[7])
if (b[4]!=0) dDenom.db4 = 30*b[2] + 210*b[4] + 6*(315*b[6]+3465*
b[8]+1)
else dDenom.db4 = 0
dDenom.db5 = 30*(b[1]+7*b[3]+63*b[5]+693*b[7])
if (b[6]!=0) dDenom.db6 = 30*(7*b[2]+63*b[4]+693*b[6]+9009*b[8]+1)
else dDenom.db6 = 0
dDenom.db7 = 210*(b[1]+9*b[3]+99*b[5]+1287*b[7])
if (b[8]!=0) dDenom.db8 = 210*(9*b[2]+99*b[4]+1287*b[6]+19305*
b[8]+1)
else dDenom.db8 = 0

if (b[1]!=0) dPoli.db1=z else dPoli.db1=0
if (b[2]!=0) dPoli.db2=z^2 else dPoli.db2=0
if (b[3]!=0) dPoli.db3=z^3 else dPoli.db3=0
if (b[4]!=0) dPoli.db4=z^4 else dPoli.db4=0
if (b[5]!=0) dPoli.db5=z^5 else dPoli.db5=0
if (b[6]!=0) dPoli.db6=z^6 else dPoli.db6=0
if (b[7]!=0) dPoli.db7=z^7 else dPoli.db7=0
if (b[8]!=0) dPoli.db8=z^8 else dPoli.db8=0

score.b1 = 2*poli*dPoli.db1/(poli^2+e0)-dDenom.db1/denom
score.b2 = 2*poli*dPoli.db2/(poli^2+e0)-dDenom.db2/denom
score.b3 = 2*poli*dPoli.db3/(poli^2+e0)-dDenom.db3/denom
score.b4 = 2*poli*dPoli.db4/(poli^2+e0)-dDenom.db4/denom
score.b5 = 2*poli*dPoli.db5/(poli^2+e0)-dDenom.db5/denom
score.b6 = 2*poli*dPoli.db6/(poli^2+e0)-dDenom.db6/denom
score.b7 = 2*poli*dPoli.db7/(poli^2+e0)-dDenom.db7/denom
score.b8 = 2*poli*dPoli.db8/(poli^2+e0)-dDenom.db8/denom

dPoli.dz = b[1] + 2*b[2]*z + 3*b[3]*z^2 + 4*b[4]*z^3 + 5*b[5]*z^4 +

```

```
6*b[6]*z^5 +7*b[7]*z^6 + 8*b[8]*z^7
```

```
if (arch.model=="egarch" | arch.model=="aparch") indicadora<-  
ifelse(z>=0,-1,1)
```

```
dSig2.da0 = numeric(n)
```

```
dSig2.da1 = numeric(n)
```

```
if (arch.model=="garch"){
```

```
if (ar==0){ c0=solution[2]; c1=solution[3]; d1=solution[4]}
```

```
else { c0=solution[3]; c1=solution[4]; d1=solution[5]}
```

```
posic.a0 = -2*c1*mean(epsilon)/(1-d1)
```

```
dSig2.da0 = .C("ds2da0garch",as.double(epsilon),as.double(c1),  
as.double(d1),as.integer(n),as.double(posic.a0),
```

```
result=double(n))$result
```

```
if (ar==1){
```

```
posic.a1 = -2*c1*mean(epsilon*c(mean(y),y[-n]))/(1-d1)
```

```
dSig2.da1 = .C("ds2da1garch",as.double(y),as.double(epsilon),  
as.double(c1),as.double(d1),as.integer(n),as.double(posic.a1),
```

```
result=double(n))$result
```

```
}
```

```
}
```

```
if (arch.model=="egarch"){
```

```
if (ar==0){ c0=solution[2]; c1=solution[3]; d1=solution[4];
```

```
m1=solution[5]}
```

```
else { c0=solution[3]; c1=solution[4]; d1=solution[5];
```

```
m1=solution[6]}
```

```
dn1 = 1-mean((sigma^2/c(mean(sigma^2),sigma[-n]^2))*(c1-0.5*  
c(mean(d1*z+m1*abs(z)),d1*z[-n]+m1*abs(z[-n]))))
```

```
posic.a0 = mean((sigma^2/c(mean(sigma),sigma[-n]))*(m1*  
c(mean(indicadora),indicadora[-n])-d1))/dn1
```

```
dSig2.da0 = .C("ds2da0egarch",as.double(sigma^2),as.double(sigma),  
as.double(z),as.double(abs(z)),as.integer(indicadora),as.double(c1),  
as.double(d1),as.double(m1),as.integer(n),as.double(posic.a0),
```

```

result=double(n))$result
if (ar==1){
posic.a1 = mean(c(mean(y),mean(y),y[-c(n-1,n)]))*
(sigma^2/c(mean(sigma),sigma[-n]))*
(m1*c(mean(indicadora),indicadora[-n])-d1))/dn1
dSig2.da1 = .C("ds2da1egarch",as.double(y),
as.double(sigma^2),as.double(sigma),
as.double(z),as.double(abs(z)),
as.integer(indicadora),as.double(c1),
as.double(d1),as.double(m1),as.integer(n),
as.double(posic.a1),result=double(n))$result
}
}
if (arch.model=="aparch"){
if (ar==0){ c0=solution[2]; c1=solution[3]; d1=solution[4];
m1=solution[5]; delta=solution[6]}
else {c0=solution[3]; c1=solution[4]; d1=solution[5];
m1=solution[6]; delta=solution[7]}
dn1 = 1-d1*mean(sigma^(2-delta)*c(mean(sigma^(delta-2)),sigma[-n]^
(delta-2)))
posic.a0 = 2*c1*mean(sigma^(2-delta)*c(mean((abs(epsilon)-m1*
epsilon)^(delta-1)),(abs(epsilon[-n])-m1*epsilon[-n])^(delta-
1))*(c(mean(indicadora),indicadora[-n])+m1)))/dn1
dSig2.da0 = .C("ds2da0aparch",as.double(sigma),as.double(epsilon),
as.double(abs(epsilon)),as.integer(indicadora),as.double(c1),
as.double(d1),as.double(m1),as.double(delta),as.integer(n),
as.double(posic.a0),result=double(n))$result
if (ar==1){
posic.a1 = 2*c1*mean(sigma^(2-delta)*c(mean((abs(epsilon)-
m1*epsilon)^(delta-1)),(abs(epsilon[-n])-m1*epsilon[-n])^(delta-
1))*(c(mean(indicadora),indicadora[-n])+m1)* c(mean(y),mean(y),
y[-c((n-1):n)])))/dn1
dSig2.da1 = .C("ds2da1aparch",as.double(y),as.double(sigma),
as.double(epsilon),as.double(abs(epsilon)),as.integer(indicadora),

```

```

as.double(c1),as.double(d1),as.double(m1),as.double(delta),
as.integer(n),as.double(posic.a1),result=double(n))$result
}
}
score.a0=(1/sigma)*(z-2*poli*dPoli.dz/(poli^2+e0))+(1/(2*sigma^2))*
(z^2-1-2*z*poli*dPoli.dz/(poli^2+e0))*dSig2.da0
if (ar==1){
score.a1=(c(mean(y),y[-n])/sigma)*(z-2*poli*dPoli.dz/(poli^2+e0))+
(1/(2*sigma^2))*(z^2-1-2*z*poli*dPoli.dz/(poli^2+e0))*dSig2.da1
}
if (ar==0) score.a1=rep(0,n)

dSig2.dc0 = numeric(n)
dSig2.dc1 = numeric(n)
dSig2.dd1 = numeric(n)
dSig2.dm1 = numeric(n)
dSig2.ddelta = numeric(n)

if (arch.model=="garch"){
posic.c0 = 1/(1-d1)
posic.c1 = mean(epsilon^2)/(1-d1)
posic.d1 = mean(sigma^2)/(1-d1)
dSig2.dc0 = .C("ds2dc0garch",as.double(d1),as.integer(n),
as.double(posic.c0),result=double(n))$result
dSig2.dc1 = .C("ds2dc1garch",as.double(epsilon^2),as.double(d1),
as.integer(n),as.double(posic.c1),result=double(n))$result
dSig2.dd1 = .C("ds2dd1garch",as.double(sigma^2),as.double(d1),
as.integer(n),as.double(posic.d1),result=double(n))$result
score.c0=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-1)*
dSig2.dc0
score.c1=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-1)*
dSig2.dc1
score.d1=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-1)*
dSig2.dd1

```

```

}

if (arch.model=="egarch"){
dn2 = 1-mean((sigma^2/c(mean(sigma^2),sigma[-n]^2))*(c1-
0.5*c(mean(d1*z+m1*abs(z)),(d1*z+m1*abs(z))[-n])))
posic.c0 = mean(sigma^2)/dn2
posic.c1 = mean(sigma^2*c(mean(log(sigma^2)),log(sigma^2)[-n]))/dn2
posic.d1 = mean(sigma^2*c(mean(z),z[-n]))/dn2
posic.m1 = mean(sigma^2*c(mean(abs(z)),abs(z)[-n]))/dn2
dSig2.dc0 = .C("ds2dc0egarch",as.double(sigma^2),
as.double(z),as.double(abs(z)),
as.double(c1),as.double(d1),as.double(m1),as.integer(n),
as.double(posic.c0),result=double(n))$result
dSig2.dc1 = .C("ds2dc1egarch",as.double(sigma^2),
as.double(log(sigma^2)),as.double(z),as.double(abs(z)),as.double(c1),
as.double(d1),as.double(m1),as.integer(n),
as.double(posic.c1),result=double(n))$result
dSig2.dd1 = .C("ds2dd1egarch",as.double(sigma^2),as.double(z),
as.double(abs(z)),as.double(c1),as.double(d1),as.double(m1),
as.integer(n),as.double(posic.d1),result=double(n))$result
dSig2.dm1 = .C("ds2dm1egarch",as.double(sigma^2),as.double(z),
as.double(abs(z)),as.double(c1),as.double(d1),as.double(m1),
as.integer(n),as.double(posic.m1),result=double(n))$result
score.c0=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-1)*
dSig2.dc0
score.c1=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-1)*
dSig2.dc1
score.d1=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-1)*
dSig2.dd1
score.m1=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-1)*
dSig2.dm1
}

if (arch.model=="aparch"){

```

```

Dnm = 1-d1*mean(sigma^(2-delta)*c(mean(sigma^(delta-2)),sigma[-n]^(
(delta-2)))
posic.c0 = 2*mean(sigma^(2-delta))/(delta*Dnm)
posic.c1 = 2*mean(sigma^(2-delta)*c(mean((abs(epsilon)-m1*epsilon)^
delta),(abs(epsilon[-n])-m1*epsilon[-n])^delta))/(delta*Dnm)
posic.d1 = 2*mean(sigma^(2-delta)*c(mean(sigma^delta),sigma[-n]^
delta))/(delta*Dnm)
posic.m1 = -2*c1*mean(sigma^(2-delta)*c(mean(epsilon*
(abs(epsilon)-m1*epsilon)^(delta-1),(epsilon*(abs(epsilon)-m1*
epsilon)^(delta-1))[-n]))/Dnm
Numd=2*mean(sigma^(2-delta)*c(mean(c1*(abs(epsilon)-m1*epsilon)^
delta*log(abs(epsilon)-m1*epsilon)+0.5*d1*sigma^delta*log(sigma^2)),
(c1*(abs(epsilon)-m1*epsilon)^delta*log(abs(epsilon)-m1*epsilon)+0.5*
d1*sigma^delta*log(sigma^2))[-n])-2*sigma^2*log(sigma^delta)/delta)
posic.delta = Numd/(delta*Dnm)
dSig2.dc0 = .C("ds2dc0aparch",as.double(sigma),as.double(d1),
as.double(delta),as.integer(n),as.double(posic.c0),
result=double(n))$result
dSig2.dc1 = .C("ds2dc1aparch",as.double(sigma),as.double(epsilon),
as.double(abs(epsilon)),as.double(d1),as.double(m1),as.double(delta),
as.integer(n),as.double(posic.c1),result=double(n))$result
dSig2.dd1 = .C("ds2dd1aparch",as.double(sigma),as.double(d1),
as.double(delta),as.integer(n),as.double(posic.d1),
result=double(n))$result
dSig2.dm1 = .C("ds2dm1aparch",as.double(sigma),as.double(epsilon),
as.double(abs(epsilon)),as.double(c1),as.double(d1),as.double(m1),
as.double(delta),as.integer(n),as.double(posic.m1),
result=double(n))$result
dSig2.ddelta = .C("ds2ddelta",as.double(sigma),as.double(epsilon),
as.double(abs(epsilon)),as.double(c1),as.double(d1),as.double(m1),
as.double(delta),as.integer(n),as.double(posic.delta),
result=double(n))$result
score.c0=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-1)*
dSig2.dc0

```

```

score.c1=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-1)*
dSig2.dc1
score.d1=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-1)*
dSig2.dd1
score.m1=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-1)*
dSig2.dm1
score.delta=(1/(2*sigma^2))*(z^2-2*z*poli*dPoli.dz/(poli^2+e0)-
1)*dSig2.ddelta
}

scores.mu = cbind(score.a0,score.a1)
colnames(scores.mu) = c("score.a0","score.a1")
scores.b = cbind(score.b1,score.b2,score.b3,score.b4,score.b5,
score.b6,score.b7,score.b8)
colnames(scores.b) = paste("score.b",1:8,sep="")
scores.b = scores.b[,-which(b==0)]

if (arch.model=="garch"){
scores=cbind(scores.mu,score.c0,score.c1,score.d1,scores.b)
}
if (arch.model=="egarch"){
scores=cbind(scores.mu,score.c0,score.c1,score.d1,score.m1,scores.b)
}
if (arch.model=="aparch"){
scores=cbind(scores.mu,score.c0,score.c1,score.d1,score.m1,
score.delta,scores.b)
}
if (ar==0) scores = scores[,-2]
return(scores)
}

```

Ahora se procede a construir la función `SNP.estim` para la estimación de las densidades SNP. Esta función tiene los siguientes argumentos: `y` es la serie; `ar` indica el número de términos autorregresivos de la ecuación de localización, puede tomar los valores de 0 ó 1; `SNPvec` indica cuáles términos del polinomio de la SNP se utilizan (se señalan

con el número 1) y cuáles se restringen a 0, por ejemplo `SNPvec=c(0,1,0,1,0,1)` significa que se utiliza un polinomio de grado 6 en la expansión SNP, y que además los coeficientes impares se restringen a 0; `arch.model` es el modelo de heteroscedasticidad condicional que se empleará para la ecuación de escala, puede ser “garch”, “egarch” o “aparch”; `spline` puede ser `TRUE` o `FALSE` e indica si se utilizará una transformación spline para los datos; si `spline=TRUE`, `c.spl` es el valor del coeficiente c para la transformación spline (ver ecuación 4.5); `par.signif` puede ser `TRUE` o `FALSE` e indica si se debe mostrar la significancia de los parámetros estimados; si `par.signif=TRUE`, `method` puede ser alguno de los métodos de optimización disponibles en la función `optim` de R, que permite obtener la significancia de los parámetros; finalmente, `startval` es un vector de parámetros iniciales para el optimizador. Esta función también genera la matriz de covarianzas asintótica (3.5) y su inversa, la matriz de ponderaciones que se utilizará posteriormente en la estimación EMM, la cual se etiqueta como `ponderaciones.EMM`.

```
SNP.estim = function(y,ar=0,SNPvec=rep(0,8),arch.model="garch",
  spline=FALSE,c.spl=0.05,par.signif=FALSE,method="BFGS",startval){
  if (ar!=0 & ar!=1) stop('ar debe ser 0 ó 1')
  if (length(SNPvec)>8) stop('La longitud de SNPvec debe ser menor
  o igual a 8')
  b.test = ifelse(SNPvec==0 | SNPvec==1,0,1)
  if (sum(b.test)!=0) stop("SNPvec está mal definido")
  if (c.spl<=0) stop('c.spl debe ser mayor a cero')

  x <- y
  n <- length(y)
  if (spline){
    x = rep(NA,n)
    for (i in (1:n)){
      if (y[i]>-c.spl & y[i]<c.spl) x[i]=y[i]
      if (y[i]<=-c.spl) x[i]=0.5*(y[i]-c.spl-log(1-y[i]-c.spl))
      if (y[i]>=c.spl) x[i]=0.5*(y[i]+c.spl+log(1+y[i]-c.spl))
    }
  }
  kz = sum(SNPvec)
```

```

SNP = function(bvec,z,sigma){
b = rep(NA,length(SNPvec))
for (i in 1:length(b)){
if (SNPvec[i]==0) b[i]=0
if (SNPvec[i]==1){b[i]=bvec[1]; bvec=bvec[-1]}
}
if (length(b)<8) b[(length(b)+1):8]=0
e0 = 0.001
denom = 1 + b[1]^2 + 6*b[1]*(b[3]+5*b[5]+35*b[7]) + 3*b[2]^2 +
2*b[2]*(15*b[4]+105*b[6]+945*b[8]+1) + 15*b[3]^2 + 210*b[3]*(b[5]+
9*b[7]) + 105*b[4]^2 + 6*b[4]*(315*b[6]+3465*b[8]+1) + 945*b[5]^2 +
20790*b[5]*b[7] + 10395*b[6]^2 + 30*b[6]*(9009*b[8]+1) + 135135*
b[7]^2 + 2027025*b[8]^2 + 210*b[8] + e0
poli = 1 + b[1]*z + b[2]*z^2 + b[3]*z^3 + b[4]*z^4 + b[5]*z^5 +
b[6]*z^6 + b[7]*z^7 + b[8]*z^8
dens = (1/(sqrt(2*pi)*sigma))*(poli^2+e0)*exp(-0.5*z^2)/denom
return(dens)
}

if (arch.model=="garch") lgt=3
if (arch.model=="egarch") lgt=4
if (arch.model=="aparch") lgt=5

logver = function(theta){
eps.sig = eps.sig.gen(theta=theta,ar=ar,y=y,x=x,arch.model=
arch.model)
epsilon = eps.sig[,1]
sigma = eps.sig[,2]
z = epsilon/sigma
if (kz>0) bvec = theta[(ar+1+lgt+1):(ar+1+lgt+kz)]
else bvec = rep(0,8)
ll = -sum(log(SNP(bvec=bvec,z=z,sigma=sigma)))
return(ll)
}

```

```

}

if (arch.model=="garch" & ar==0){
lower = c(-1e5,1e-10,1e-10,1e-10,rep(-20,kz))
upper = c(1e5,1e5,1-1e-10,1-1e-10,rep(20,kz))
}
if (arch.model=="garch" & ar==1){
lower = c(-1e5,-1+1e-10,1e-10,1e-10,1e-10,rep(-20,kz))
upper = c(1e5,1-1e-10,1e5,1-1e-10,1-1e-10,rep(20,kz))
}
if (arch.model=="egarch" & ar==0){
lower = c(-1e5,-1e2,-1+1e-10,-1+1e-10,-1+1e-10,rep(-20,kz))
upper = c(1e5,1e2,1-1e-10,1-1e-10,1-1e-10,rep(20,kz))
}
if (arch.model=="egarch" & ar==1){
lower = c(-1e5,-1+1e-10,-1e2,-1+1e-10,-1+1e-10,-1+1e-10,rep(-20,kz))
upper = c(1e5,1-1e-10,1e2,1-1e-10,1-1e-10,1-1e-10,rep(20,kz))
}
if (arch.model=="aparch" & ar==0){
lower = c(-1e5,1e-10,1e-10,1e-10,-1+1e-10,1e-10,rep(-20,kz))
upper = c(1e5,1e5,1-1e-10,1-1e-10,1-1e-10,4,rep(20,kz))
}
if (arch.model=="aparch" & ar==1){
lower = c(-1e5,-1+1e-10,1e-10,1e-10,1e-10,-1+1e-10,1e-10,rep(-20,kz))
upper = c(1e5,1-1e-10,1e5,1-1e-10,1-1e-10,1-1e-10,4,rep(20,kz))
}

if (par.signif){
sol = optim(par=startval,fn=logver,lower=lower,upper=upper,
method=method,hessian=TRUE)
hess = sol$hessian
std.err = sqrt(diag(solve(hess)))
tval = sol$par/std.err
valp = ifelse(tval>0,(pnorm(tval,lower.tail=F)+pnorm(-tval,

```

```

lower.tail=T)),(pnorm(tval,lower.tail=T)+pnorm(-tval,lower.tail=F)))
logveros = -sol$value
solution = sol$par
}
else {
sol = nlminb(start=startval,objective=logver,lower=lower,
upper=upper,control=list(eval.max=2000,iter.max=1500))
solution = sol$par
logveros = -sol$objective
}

bic = -2*logveros + log(n)*length(solution)
mu.par = solution[1:(ar+1)]
names(mu.par) = paste("a",0:ar,sep="")

if (arch.model=="garch"){
sig.par = solution[(ar+2):(ar+4)]
names(sig.par) = c("c0","c1","d1")
}
if (arch.model=="egarch"){
sig.par = solution[(ar+2):(ar+5)]
names(sig.par) = c("c0","c1","d1","m1")
}
if (arch.model=="aparch"){
sig.par = solution[(ar+2):(ar+6)]
names(sig.par) = c("c0","c1","d1","m1","delta")
}

lgth = length(mu.par)+length(sig.par)
if(kz>0){
b.par = solution[(lgth+1):length(solution)]
names.b = paste("b",1:length(SNPvec),sep="")
names(b.par) = names.b[-which(c(SNPvec,0)==0)]
}

```

```

else b.par = rep(0,8)

if (kz>0) names(solution) = c(names(mu.par),names(sig.par),
names(b.par))
else names(solution) = c(names(mu.par),names(sig.par))

epsdes = eps.sig.gen(theta=solution,ar=ar,y=y,x=x,arch.model=
arch.model)
epsilon = epsdes[,1]
sigma = epsdes[,2]
z = epsilon/sigma
scores = SNP.score.gen(solution=solution,SNPvec=SNPvec,bvec=b.par,
z=z,arch.model=arch.model,epsilon=epsilon,sigma=sigma,y=y,ar=ar)

suma = scores[1,]%*%t(scores[1,])
for (i in (2:n)){
aux = scores[i,]%*%t(scores[i,])
suma = suma + aux
}
matrizI = suma/n
inversa = solve(matrizI)

if (par.signif){
signific = rbind(solution,std.err,tval,valp)
colnames(signific) = names(solution)
rownames(signific) = c("Coefic.," "Std.Err","t calc.," "Valor-P")
}

adj = c(logveros,bic)
names(adj) = c("LogLik","BIC")

if (par.signif){
final = list(coeficientes=solution,significancia=signific,
Ajuste=adj,ponderaciones.EMM=inversa,prom.scores=matrizI)

```

```

}
else {
final = list(coeficientes=solution,Ajuste=adj,
ponderaciones.EMM=inversa,prom.scores=matrizI)
}
return(final)
}

```

La siguiente es una función auxiliar que calcula los scores promedio (ecuación 3.3), los cuales se utilizan en el criterio EMM (ecuación 3.4).

```

SNP.scores = function(QMLEs,y,SNPvec=rep(0,8),ar=0,
arch.model="garch",spline=F,c.spl=0.1){
x = y
n = length(y)
if (spline){
x = rep(NA,n)
for (i in (1:n)){
if (y[i]>-c.spl & y[i]<c.spl) x[i]=y[i]
if (y[i]<=-c.spl) x[i]=0.5*(y[i]-c.spl-log(1-y[i]-c.spl))
if (y[i]>=c.spl) x[i]=0.5*(y[i]+c.spl+log(1+y[i]-c.spl))
}
}
epsdes = eps.sig.gen(theta=QMLEs,ar=ar,y=y,x=x,arch.model=
arch.model)
epsilon = epsdes[,1]
sigma = epsdes[,2]
z = epsilon/sigma
kz = sum(SNPvec)
if (kz>0) bvec = QMLEs[(length(QMLEs)-kz+1):length(QMLEs)]
else b=rep(0,8)
scores = SNP.score.gen(solution=QMLEs,SNPvec=SNPvec,bvec=bvec,z=z,
arch.model=arch.model,epsilon=epsilon,sigma=sigma,y=y,ar=ar)
scores.prom = colMeans(scores)
return(scores.prom)
}

```

}

Por último, la función `EMM.estim` permite estimar un modelo de volatilidad estocástica mediante EMM, para los casos en que se asume que ε_t sigue una distribución normal o una distribución t con 3 grados de libertad. El método recurre a una función que genera simulaciones del modelo SV respectivo, `SVsim`, con el propósito de obtener los scores promedio en los que se basa el criterio EMM. Los argumentos de `EMM.estim` son los siguientes: `y` es la serie, `Nsim` es el número de simulaciones para calcular los scores promedio, `signif` puede ser `TRUE` o `FALSE` e indica si se debe mostrar la significancia del modelo SV y de los coeficientes estimados, `cond.dist` indica si se estima un modelo SV con ε_t siguiendo una distribución normal estándar (“normal”) o una t estandarizada con 3 grados de libertad (“t”), `SV.start` es un vector de coeficientes iniciales $(\alpha, \phi, \sigma_\eta, \rho)$ para el optimizador. Los argumentos de `mod.aux` son los que corresponden al modelo auxiliar SNP y son los mismos de la función `SNP.estim`, descrita anteriormente.

```
EMM.estim = function(y,Nsim=80000,signif=FALSE,cond.dist="normal",
SV.start,mod.aux=list(ar=0,SNPvec=rep(0,8),arch.model="garch",
spline=F,c.spl=0.05,startval)){
ar = mod.aux$ar
SNPvec = mod.aux$SNPvec
arch.model = mod.aux$arch.model
spline = mod.aux$spline
c.spl = mod.aux$c.spl
startval = mod.aux$startval

if (cond.dist=="normal") epsilon = rnorm(Nsim+100,0,1)
if (cond.dist=="t") epsilon = sqrt(1/3)*rt(Nsim+100,df=3)
e = rnorm(Nsim+100,0,1)
SNP = SNP.estim(y=y,ar=ar,SNPvec=SNPvec,arch.model=arch.model,
spline=spline,c.spl=c.spl,par.signif=FALSE,startval=startval)
coef.aux = SNP$coeficientes
ponderac = SNP$ponderaciones.EMM

criterio = function(lambda){
```

```

Xsim = SVsim(n=Nsim,coef=lambda,input.random=TRUE,
epsilon=epsilon,e=e)[,1]
score.prom = SNP.scores(QMLEs=coef.aux,y=Xsim,SNPvec=SNPvec,ar=ar,
arch.model=arch.model,spline=spline,c.spl=c.spl)
crit = t(score.prom)%*%ponderac%*%score.prom
return(crit)
}
upper = c(0,1,5,1)
lower = c(-20,0.0001,0.0001,-1)
soluc = nlminb(start=SV.start,objective=criterio,upper=upper,
lower=lower,control=list(eval.max=150,iter.max=100))
resultado = soluc$par
names(resultado) = c("alfa","phi","sigma.eta","rho")
salida = list(coeficientes=resultado)

if (signif){
n1 = length(y)
Xoptim = SVsim(n=Nsim,coef=resultado,input.random=TRUE,
epsilon=epsilon,e=e)[,1]
sc.pr.est = SNP.scores(QMLEs=coef.aux,y=Xoptim,SNPvec=SNPvec,ar=ar,
arch.model=arch.model,spline=spline,c.spl=c.spl)
chi.cuad = n1*t(sc.pr.est)%*%ponderac%*%sc.pr.est
glib = length(coef.aux)-length(resultado)
SVvalp = pchisq(chi.cuad,df=glib,lower.tail=FALSE)
modelsignif = c(chi.cuad,round(glib,0),SVvalp)
names(modelsignif) = c("Chi-cuad","g.l.,""Valor-P")

G = function(lambda){
Xoptim = SVsim(n=Nsim,coef=lambda,input.random=TRUE,
epsilon=epsilon,e=e)[,1]
SC = SNP.scores(QMLEs=coef.aux,y=Xoptim,SNPvec=SNPvec,ar=ar,
arch.model=arch.model,spline=spline,c.spl=c.spl)
return(SC)
}

```



```

M = numDeriv::jacobian(func=G,x=resultado)
MIM = solve(t(M)%*%ponderac%*%M)
vcov = (1/n1)*MIM
std.err = sqrt(diag(vcov))
tval = resultado/std.err
valp = ifelse(tval>0,(pnorm(tval,lower.tail=F)+pnorm(-tval,
lower.tail=T)),(pnorm(tval,lower.tail=T)+pnorm(-tval,lower.tail=F)))
coefic = matrix(nrow=4,ncol=length(resultado),rbind(resultado,
std.err,tval,valp))
colnames(coefic)=names(resultado)
rownames(coefic)=c("Coefic. ","Std.Err", "t calc.", "Valor-P")
Imat = SNP$prom.scores
Smat = sqrt(diag(Imat-M%*%MIM%*%t(M)))
cuasi.t = sqrt(n1)*sc.pr.est/Smat
names(cuasi.t) = names(coef.aux)
salida = list(SV.model.signif=modelsignif,coeficientes=coefic,
cuasi.t.ratios=t(cuasi.t))
}
return(salida)
}

```

Ejemplo

Suponga que se tiene una serie de retornos de acciones, llamada y , a la que queremos ajustar un modelo SV log-normal. Lo primero que hacemos es cargar la librería DLL “funcionesC.dll” y correr todas las funciones descritas anteriormente. El siguiente paso es encontrar un modelo auxiliar que describa adecuadamente los datos. La estimación del modelo auxiliar se hace recurriendo a la función `SNP.estim`. Por ejemplo, un modelo auxiliar AR(1)-APARCH(1,1)-SNP(4) con transformación spline se estima de la siguiente manera:

```

iniciales=c(-0.0006,-0.5, 0.02,0.7, 0.15,0.5,1.5,rep(0.01,4))
modaux = SNP.estim(y=y,ar=1,SNPvec=rep(1,4),arch.model="aparch",
spline=T,c.spl=0.03,par.signif=T,method="BFGS",startval=iniciales)

```

Aquí se debe tener en cuenta que es necesario probar diferentes valores para el vector

de parámetros iniciales, con el objetivo de evitar caer en óptimos locales. Se escogen los parámetros cuya estimación alcanza la máxima log-verosimilitud, la cual es reportada por `SNP.estim`. Luego de elegir el modelo auxiliar adecuado, probando modelos con `AR(0)`, con términos `EGARCH` y variando el número de términos de la expansión `SNP`, y luego seleccionando según el criterio `BIC` y la significancia de los parámetros, se procede a la estimación `EMM`. Supongamos que el modelo auxiliar `AR(1)-APARCH(1,1)-SNP(4)` es adecuado, entonces la estimación `EMM` se ejecuta escribiendo:

```
inicialSV = c(-7.810,0.95,0.32730,-0.7)
SVmod = EMM.estim(y=y,Nsim=80000,signif=T,cond.dist="normal",
SV.start=inicialSV,mod.aux=list(ar=1,SNPvec=rep(1,4),arch.model=
"aparch",spline=T,c.spl=0.03,startval=iniciales)))
```

En este paso también se debe evaluar la estimación arrancando desde diferentes valores iniciales del vector de parámetros del modelo `SV`.