

# Sobre la estimación del parámetro de grosor de cola de la distribución de variación regular en procesos GARCH

## Sensibilidad de las estimaciones a la distribución y propiedades de las innovaciones, errores de especificación de la varianza condicional, estimación de los parámetros GARCH y estabilidad bajo agregación

Afonso Rodríguez, Julio Angel  
UNIVERSIDAD DE LA LAGUNA

Departamento de Economía de las Instituciones, Estadística Económica y Econometría  
Instituto Universitario de Desarrollo Regional (IUDR)  
Facultad de CC. Económicas y Empresariales  
Camino La Hornera, s/n. Campus de Guajara. 38202 LA LAGUNA  
Tenerife. Islas Canarias  
Tfno.: 922.317041, Fax: 922.317042  
email: jafonsor@ull.es

### RESUMEN:

Los modelos de heterocedasticidad condicionada dinámica tipo GARCH representan en la actualidad una de las herramientas econométricas más importantes en la descripción de los procesos de rendimientos de activos financieros. Estos modelos permiten capturar razonablemente las principales características de estas series de rendimientos: distribuciones no normales con colas pesadas, agrupamiento de la volatilidad, etc. Habitualmente la estimación se realiza empleando máxima verosimilitud bajo distintas distribuciones de las innovaciones (gaussiana, t-Student, GED, mezcla de normales,...) y la inferencia se realiza comúnmente bajo el supuesto de consistencia y distribución asintótica normal, independientemente de la distribución utilizada. Sin embargo, Berkes y Horváth (2003) indican cierta cautela en relación con el supuesto de consistencia puesto que conectan ésta con el número de momentos finitos en la distribución de las innovaciones, mientras que Mikosch y Stàricà (2000) establecen la que las distribuciones de dimensión finita de la solución estrictamente estacionaria de un proceso GARCH(1,1) es de variación regular con parámetro de grosor de cola  $\kappa > 0$ , de forma que las distribuciones marginales caen en el dominio de atracción de una ley estable tipo Lévy-Pareto. Basrak, Davis y Mikosch (2002) extienden estos resultados al caso GARCH( $p,q$ ). A través del cumplimiento de una condición de existencia de momento, Stàricà y Pictet (1997), Berkes, Horváth y Kokoszka (2003) e Iglesias y Linton (2007) proponen un estimador de  $\kappa$ . Motivado por estos recientes descubrimientos, en este trabajo revisamos los principales resultados relativos a las características de la distribución marginal de colas gruesas de un proceso GARCH(1,1) y las implicaciones sobre la inferencia en relación con la posibilidad de estimar el parámetro  $\kappa$  que gobierna las propiedades fundamentales de esta distribución y el comportamiento de los estadísticos habituales basados en un proceso GARCH. Examinamos las principales características del estimador de  $\kappa$  propuesto por Starica y Pictet (1997) e Iglesias y Linton (2007) con el objeto de una mejor interpretación de los resultados que se obtienen en aplicaciones con datos reales. Esta metodología se aplica también a datos reales de tipos de cambio (\$/euro, yen/\$ y \$/libra esterlina).

**Palabras Clave:** *Procesos GARCH, leyes de escalamiento o leyes de potencia, parámetro de índice de cola o de estabilidad, auto-similaridad, agregación temporal, tipos de cambio.*

**Clasificación Código JEL:** *C22, C15, C52, G12*

**ÁREA TEMÁTICA:** *MERCADOS MONETARIOS Y FINANCIEROS  
MÉTODOS CUANTITATIVOS EN ECONOMÍA*

# **Sobre la estimación del parámetro de grosor de cola de la distribución de variación regular en procesos GARCH**

**Sensibilidad de las estimaciones a la distribución y propiedades de las innovaciones, errores de especificación de la varianza condicional, estimación de los parámetros GARCH y estabilidad bajo agregación**

## **1. Introducción**

Es ampliamente conocido que los modelos GARCH permiten describir algunas de las principales características empíricas de las series temporales financieras o económicas de alta frecuencia como exceso de curtosis (probabilidades significativas de observaciones extremas), varianza posiblemente infinita, agrupamiento de la volatilidad, ... Los resultados recientes de Mikosch y Starica (2000) y Basrak, Davis y Mikosch (2002) proporcionan una justificación formal a esta capacidad de replicar estas propiedades al caracterizar la naturaleza de variación regular con colas gruesas de las distribuciones de dimensión finita de los procesos GARCH así como las distribuciones marginales tipo Pareto con estructura de ley de potencia básicamente caracterizadas por un parámetro,  $\kappa$ , que gobierna la forma y principales características de estas distribuciones. Además, en relación con los procedimientos de inferencia en modelos GARCH, el valor de  $\kappa$  determina la diferencia entre la aplicabilidad de la ley de los grandes números y el teorema central del límite, o bien la convergencia en ley y distribución a distribuciones estables como muestran Mikosch y Starica (2000), Hall y Yao (2003) y Mikosch y Straumann (2006).

En este trabajo revisamos algunos de los principales resultados establecidos en estas aportaciones y nos centramos en dos propuestas para la estimación del parámetro de grosor de cola de la distribución  $\kappa$ , que a diferencia de los estimadores tipo Hill (1975) basados en un cierto número de estadísticos de orden superiores de la distribución condicional, hacen uso de la estructura particular de la distribución incondicional de los procesos GARCH. Estas son las propuestas de Starica y Pictet (1997) y de Iglesias y Linton (2007). En particular, realizamos un pequeño estudio sobre el comportamiento y principales características de estos estimadores, especialmente dirigido a la aplicación práctica de los mismos.

## 2. GARCH(1,1): representación, estacionariedad estricta y distribución marginal de variación regular

Para describir las características básicas de muchas series temporales de rendimientos de activos financieros, modelos de la forma

$$r_t = z_t \sqrt{h_t(\boldsymbol{\theta})} \quad t=1, \dots, T \quad (2.1)$$

se han convertido en muy populares. Bajo el supuesto de que el proceso de innovaciones,  $z_t$ , tiene media incondicional nula, varianza unitaria y es independiente de  $r_{t-i}$ ,  $i > 0$ ,  $h_t(\boldsymbol{\theta})$  es la varianza condicional o volatilidad estocástica de  $r_t$ . De entre todas las parametrizaciones existentes de  $h_t(\boldsymbol{\theta})$ , en este trabajo nos ocuparemos del modelo GARCH(1,1) debido a Bollerslev (1986), dado por

$$h_t(\boldsymbol{\theta}) = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 h_{t-1}(\boldsymbol{\theta}), \quad \boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1)$$

donde  $\alpha_0 > 0$ ,  $\alpha_1 > 0$  y  $\beta_1 > 0$  para asegurar la positividad de la varianza condicional. Puesto que en lo que sigue en este trabajo, los resultados parecen ser aplicables tanto en el caso de que  $h_t(\boldsymbol{\theta})$  describa la varianza condicional del proceso directamente observable de rendimientos  $r_t$  como en el caso de que  $h_t(\boldsymbol{\theta})$  sea la varianza condicional del término de error de un modelo de regresión o tipo ARMA para los niveles de la serie,  $r_t = \mu_t + \varepsilon_t$ ,  $\varepsilon_t = z_t \sqrt{h_t(\boldsymbol{\theta})}$ , consideraremos en general esta última representación, con  $h_t(\boldsymbol{\theta})$  de la forma

$$h_t(\boldsymbol{\theta}) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}(\boldsymbol{\theta}) \quad (2.2)$$

Si las innovaciones en (2.1) vienen dadas por un proceso *i.i.d.*, el proceso GARCH se denomina GARCH fuerte, mientras que si  $z_t$  viene dado por un proceso estrictamente estacionario y ergódico, se denomina GARCH semi-fuerte, de forma que en este caso la dinámica del comportamiento de  $\varepsilon_t$  no viene dado únicamente por la evolución de la varianza condicional, puesto que el proceso  $z_t$  puede presentar correlación serial. El proceso de varianza condicional GARCH(1,1) es débilmente estacionario si se cumple que  $0 < \alpha_1 + \beta_1 < 1$ , de forma que la varianza y la estructura de covarianzas y correlaciones incondicionales está definida y es finita. Sin embargo, el análisis que se llevará a cabo en este trabajo requiere la estacionariedad estricta del proceso GARCH(1,1) que, en general, determina un espacio paramétrico admisible más amplio. Basrak, Davis y Mikosch (2002) establecen el resultado de que la distribución marginal estrictamente estacionaria de un proceso GARCH es regularmente variable, es de cir,

que las colas de esta distribución son de tipo Pareto y, por lo tanto, de colas pesadas. La variación regular de la distribución conjunta de  $(\varepsilon_t^2, h_t(\boldsymbol{\theta}))$  proporciona evidencia sobre las propiedades de los momentos del proceso así como sobre la estructura de dependencia entre observaciones próximas (propiedades mixing) (ver, por ejemplo, Carrasco y Chen (2002) y Lindner (2007)), fundamental para establecer las propiedades de la distribución asintótica de los estimadores de los parámetros y de estadísticos como las funciones de autocovarianzas y autocorrelaciones muestrales de un proceso GARCH. La variación regular de estas distribuciones es una consecuencia del hecho de que los cuadrados de un proceso GARCH estacionario pueden ser incrustados en una ecuación de recurrencia estocástica (SRE: *stochastic recurrence equation*). Para este tipo de ecuaciones recursivas existe una teoría avanzada debida a Kesten (1973) que proporciona condiciones para la existencia de una única solución estacionaria al sistema y que describe el comportamiento de las colas de la distribución de la solución estacionaria. Liu (2006) extiende estos resultados al caso de la familia de procesos GARCH de He y Teräsvirta (1999) que incluye como casos particulares muchas de las especificaciones tipo GARCH(1,1) habitualmente utilizadas en la práctica.

## 2.1 Representación SRE: unidimensional y bidimensional

### Representación SRE unidimensional:

Haciendo uso de la representación

$$\varepsilon_t^2 = z_t^2 \cdot h_t(\boldsymbol{\theta}) \quad (2.3)$$

para el cuadrado del proceso GARCH, la ecuación GARCH(1,1) puede escribirse como

$$h_t(\boldsymbol{\theta}) = \alpha_0 + (\alpha_1 z_{t-1}^2 + \beta_1) h_{t-1}(\boldsymbol{\theta}) = \alpha_0 + A_{t-1}(\boldsymbol{\theta}) h_{t-1}(\boldsymbol{\theta}) \quad (2.4)$$

Mediante sustituciones recursivas y considerando que el proceso comienza indefinidamente lejos, se tiene que, haciendo  $\prod_{j=1,0}^{\infty} A_{t-j}(\boldsymbol{\theta}) = 1$ ,

$$h_t(\boldsymbol{\theta}) = \alpha_0 \sum_{i=1}^{\infty} \prod_{j=1}^{i-1} A_{t-j}(\boldsymbol{\theta}) + h_{-\infty}(\boldsymbol{\theta}) \prod_{j=1}^{\infty} A_{t-j}(\boldsymbol{\theta}) \rightarrow h_{\infty}(\boldsymbol{\theta}) = \alpha_0 \sum_{i=1}^{\infty} \prod_{j=1}^{i-1} A_{t-j}(\boldsymbol{\theta}) \text{ c.s.} \quad (2.5)$$

donde  $h_{\infty}(\boldsymbol{\theta})$  es la única solución estacionaria en distribución. Asumiendo que las innovaciones del proceso son *i.i.d.*, por aplicación de la ley fuerte de los grandes números, la secuencia  $\prod_{j=1}^m A_{t-j}(\boldsymbol{\theta}) \rightarrow 0$  c.s. cuando  $m \rightarrow \infty$  bajo los supuestos  $\ln^+(A_0(\boldsymbol{\theta})) < \infty$  y  $E[\ln A_0(\boldsymbol{\theta})] < 0$ . Por otro lado, bajo los mismos supuestos, la

secuencia  $\sum_{i=1}^m \prod_{j=1}^{i-1} A_{t-j}(\boldsymbol{\theta})$  converge a un límite finito y, por tanto,  $h_{\infty}(\boldsymbol{\theta})$  está bien definido.

### Representación SRE bidimensional:

La representación bidimensional del proceso conjunto GARCH(1,1),  $(\varepsilon_t^2, h_t(\boldsymbol{\theta})) = (z_t^2 h_t(\boldsymbol{\theta}), h_t(\boldsymbol{\theta}))$  de la forma

$$\begin{pmatrix} \varepsilon_t^2 \\ h_t(\boldsymbol{\theta}) \end{pmatrix} = \mathbf{A}_t(\boldsymbol{\theta}) \begin{pmatrix} \varepsilon_{t-1}^2 \\ h_{t-1}(\boldsymbol{\theta}) \end{pmatrix} + \mathbf{B}_t(\boldsymbol{\theta}), \quad \mathbf{A}_t(\boldsymbol{\theta}) = \begin{pmatrix} z_t^2 \alpha_1 & z_t^2 \beta_1 \\ \alpha_1 & \beta_1 \end{pmatrix}, \quad \mathbf{B}_t(\boldsymbol{\theta}) = \alpha_0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (2.6)$$

permite encontrar la solución estacionaria en distribución del proceso bidimensional  $(\varepsilon_t^2, h_t(\boldsymbol{\theta}))$ , así como la distribución marginal de cada uno de los componentes, haciendo uso de los resultados generales de Kesten (1973). Alternativamente a (2.6), se puede escribir de forma equivalente como

$$\begin{pmatrix} \varepsilon_t^2 \\ h_{t+1}(\boldsymbol{\theta}) \end{pmatrix} = \mathbf{A}_t(\boldsymbol{\theta}) \begin{pmatrix} \varepsilon_{t-1}^2 \\ h_t(\boldsymbol{\theta}) \end{pmatrix} + \mathbf{B}_t(\boldsymbol{\theta}), \quad \mathbf{A}_t(\boldsymbol{\theta}) = \begin{pmatrix} z_t^2 & 0 \\ \beta_1 + z_t^2 \alpha_1 & 0 \end{pmatrix}, \quad \mathbf{B}_t(\boldsymbol{\theta}) = \begin{pmatrix} 0 \\ \alpha_0 \end{pmatrix}$$

que proporciona los mismos resultados. Llamando  $\mathbf{Y}_t = (\varepsilon_t^2, h_{t+1}(\boldsymbol{\theta}))$ , se tiene la representación recursiva siguiente:

$$\mathbf{Y}_t = \prod_{j=1}^n A_{t-i+1}(\boldsymbol{\theta}) \mathbf{Y}_{t-n} + \mathbf{B}_t(\boldsymbol{\theta}) + \sum_{i=1}^{n-1} \mathbf{B}_{t-i}(\boldsymbol{\theta}) \prod_{j=1}^i A_{t-j+1}(\boldsymbol{\theta}) \quad (2.7)$$

## 2.2 Condición de estacionariedad estricta

A partir de la representación (2.7), el resultado fundamental para establecer la estacionariedad estricta y la representación causal única del proceso GARCH conjunto, viene dado por el valor negativo del denominado exponente superior de Lyapunov, definido como

$$\gamma(\boldsymbol{\theta}) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \|\mathbf{A}_1(\boldsymbol{\theta}) \cdots \mathbf{A}_n(\boldsymbol{\theta})\| = \inf_{n \in \mathbb{N}} \left\{ \frac{1}{n} E \ln \|\mathbf{A}_1(\boldsymbol{\theta}) \cdots \mathbf{A}_n(\boldsymbol{\theta})\| \right\} \text{ c.s.}$$

donde para una matriz cuadrada  $\mathbf{A}$ ,  $\|\mathbf{A}\| = \sup_{|\mathbf{x}|=1} |\mathbf{x} \cdot \mathbf{A}|$  (norma supremo). Dada la estructura de las matrices  $\mathbf{A}_t(\boldsymbol{\theta})$  en cualquiera de las representaciones anteriores, se verifica la relación  $\mathbf{A}_t(\boldsymbol{\theta}) \cdot \mathbf{A}_{t+1}(\boldsymbol{\theta}) = (\alpha_1 z_{t+1}^2 + \beta_1) \cdot \mathbf{A}_t(\boldsymbol{\theta})$   $i=1, \dots, n-1$ , de forma que en el

caso GARCH(1,1),  $\gamma(\boldsymbol{\theta})$  toma la forma simple  $\gamma(\boldsymbol{\theta}) = \frac{1}{n} \sum_{i=1}^n E[\ln A_i(\boldsymbol{\theta})]$ .

**Proceso GARCH(1,1) fuerte:** Puesto que en este caso, las innovaciones representan un

proceso *i.i.d.*, es inmediato comprobar que

$$\gamma(\boldsymbol{\theta}) = E[\ln A_0(\boldsymbol{\theta})] = E[\ln(\alpha_1 z_0^2 + \beta_1)]$$

**Proceso GARCH(1,1) semi-fuerte:** En este caso se asume que las innovaciones,  $z_t$ , son un proceso estrictamente estacionario y ergódico con media nula y varianza unitaria, presentando, por tanto, un cierto grado de dependencia serial o correlación. Así, la secuencia de funciones  $A_{t-j}(\boldsymbol{\theta}) = \alpha_1 z_{t-j}^2(\boldsymbol{\theta}) + \beta_1$  forma un proceso estrictamente estacionario y ergódico, de forma que por la ley fuerte de los grandes números para estos procesos se verifica que la condición de estacionariedad estricta y convergencia en distribución a un límite no degenerado viene dado por la misma condición que en el caso GARCH fuerte,  $\gamma(\boldsymbol{\theta}) = E[\ln A_0(\boldsymbol{\theta})] < 0$  (ver Dahl e Iglesias (2007)).

Sólo en algunos pocos casos particulares se puede obtener una expresión de  $\gamma(\boldsymbol{\theta})$  que puede ser evaluada analíticamente para valores dados de los parámetros GARCH. En el caso ARCH(1) ( $\beta_1 = 0$ ), se tiene que  $\gamma(\boldsymbol{\theta}) = \ln(\alpha_1) + 2E[\ln z_0^2]$ . En el caso GARCH(1,1), Nelson (1990) (Teorema 6) proporciona expresiones explícitas para  $\gamma(\boldsymbol{\theta})$  en el caso de innovaciones  $N(0,1)$  y estándar Cauchy (con función de densidad  $g(z) = 1/(\pi(1+z^2))$ ) de la forma:

$N(0,1)$ :

$$N(0,1) : \gamma(\boldsymbol{\theta}) = \ln(2\alpha_1) + \psi(1/2) + (2\pi\beta_1 / \alpha_1)^{1/2} \Phi\left(\frac{1}{2}, 1.5; \beta_1 / 2\alpha_1\right) - (\beta_1 / \alpha_1) {}_2F_2(1, 1; 2, 1.5, \beta_1 / 2\alpha_1)$$

donde  $\psi(\cdot)$  es la función psi de Euler,  $\Phi(\cdot)$  es una función hipergeométrica confluyente y  ${}_2F_2(\cdot)$  es una función hipergeométrica generalizada.

$$\text{Estándar Cauchy: } \gamma(\boldsymbol{\theta}) = 2\ln(\beta_1^{1/2} + \alpha_1^{1/2})$$

En cualquier otro caso, la evaluación de  $\gamma(\boldsymbol{\theta})$  requiere, en general, integración numérica o evaluación mediante simulación. Si además de verificarse la condición de negatividad del exponente de Lyapunov,  $\gamma(\boldsymbol{\theta}) < 0$ , se tiene que  $E \ln^+ \|A_0(\boldsymbol{\theta})\| < \infty$  y  $E \ln^+ |B_0(\boldsymbol{\theta})| < \infty$  entonces la serie definida por

$$Y_t = B_t(\boldsymbol{\theta}) + \sum_{i=1}^{\infty} B_{t-i}(\boldsymbol{\theta}) \prod_{j=1}^i A_{t-j+1}(\boldsymbol{\theta}) \quad (2.8)$$

converge casi seguro a un límite no degenerado y es la única solución causal estrictamente estacionaria de (2.6). La condición de estacionariedad estricta es más débil que la condición de estacionariedad en covarianza, de forma que en el caso GARCH(1,1), la restricción  $0 < \alpha_1 + \beta_1 < 1$  que determina esta última implica aquélla.

### 2.3 Distribución marginal de variación regular

Bajo un conjunto de condiciones adicionales a las establecidas en el apartado 2.1 (ver, por ejemplo, Basrak *et.al.* (2002) o Starica y Pictet (1997)), las distribuciones de dimensión finita de la solución estrictamente estacionaria de  $Y_t$  son regularmente variables, donde la distribución marginal del proceso GARCH estacionario exhibe un comportamiento de ley de potencia tipo Pareto con índice  $\kappa$ , de la forma

$$P(v_t(\boldsymbol{\theta}) > x) \sim c_0 x^{-2\kappa}, \quad x \rightarrow \infty, \quad v_t(\boldsymbol{\theta}) = \sqrt{h_t(\boldsymbol{\theta})}$$

y

$$P(|\varepsilon_t| > x) = P(v_t(\boldsymbol{\theta}) | z_t| > x) \sim E(|z_t|^{2\kappa}) P(v_t(\boldsymbol{\theta}) > x)$$

con

$$c_0 = \frac{E[(\alpha_0 + A_0(\boldsymbol{\theta})h_0(\boldsymbol{\theta}))^\kappa - (A_0(\boldsymbol{\theta})h_0(\boldsymbol{\theta}))^\kappa]}{\kappa E[A_0^\kappa(\boldsymbol{\theta}) \ln A_0(\boldsymbol{\theta})]} > 0$$

y donde  $\kappa \in (0, \kappa_0]$ , un entero no par, viene dado por la única solución a la ecuación

$$0 = \lim_{n \rightarrow \infty} \frac{1}{n} \ln E \| A_1(\boldsymbol{\theta}) \cdots A_n(\boldsymbol{\theta}) \|^\kappa$$

que en el caso GARCH(1,1) viene dado simplemente por la ecuación equivalente

$$E[(\alpha_1 z_0^2 + \beta_1)^\kappa] = 1 \tag{2.9}$$

Así,  $\kappa$  determina el índice de la distribución marginal de variación regular a infinito de la solución causal estrictamente estacionaria de  $(\varepsilon_t^2, h_t(\boldsymbol{\theta}))$  y, como consecuencia, determina que  $\kappa_1 = (\kappa/2)$  es el índice de la distribución regularmente variable a  $\infty$  y  $-\infty$  del proceso  $(\varepsilon_t, v_t(\boldsymbol{\theta}))$ . Además,  $\kappa$  determina el máximo momento (fraccional) finito de la distribución de  $(\varepsilon_t^2, h_t(\boldsymbol{\theta}))$  en (2.8), denominado en Berkes, Horváth y Kokoszka (2003) exponente de momento máximo. Así, si  $m$  es el mayor entero positivo estrictamente menor que  $\kappa$ , entonces  $m$  es el orden del máximo momento finito de la distribución incondicional regularmente variable de  $(\varepsilon_t^2, h_t(\boldsymbol{\theta}))$ .

Puesto que la ecuación (2.9) depende de  $\kappa$ ,  $\boldsymbol{\theta}$  y de la distribución asumida de las innovaciones que puede depender de algún parámetro (o conjunto de parámetros) adicional,  $\eta$ , se puede escribir de forma alternativa como

$$\psi(\kappa, \boldsymbol{\theta} | \eta) = E[(\alpha_1 z_0^2 + \beta_1)^\kappa] - 1 = \gamma_\kappa(\boldsymbol{\theta} | \eta) - 1 = 0 \tag{2.10}$$

Empleando un desarrollo en serie binomial de  $\gamma_\kappa(\boldsymbol{\theta} | \eta)$  en (2.10) se tiene que

$$\gamma_{\kappa}(\boldsymbol{\theta} | \eta) = \beta_1^{\kappa} \sum_{k=0}^{\infty} \binom{\kappa}{k} (\alpha_1 / \beta_1)^k E(z_0^{2k})$$

de forma que tanto los valores de los coeficientes GARCH(1,1) como las características de la distribución de las innovaciones influirán en la determinación del valor numérico de  $\kappa$  como solución de (2.10). Cuánto menos gruesas sean las colas de la distribución de las innovaciones (bajo simetría), mayor será en general el número de términos sumados hasta alcanzar el valor unitario en la función  $\gamma_{\kappa}(\boldsymbol{\theta} | \eta)$  y, por tanto, mayor será el valor de  $\kappa$ . Sin embargo, incluso para distribuciones no excesivamente leptocúrticas de las innovaciones, bajo determinadas combinaciones de  $(\alpha_1, \beta_1)$  el resultado puede ser un valor relativamente bajo de  $\kappa$ , indicando así el carácter leptocúrtico de la distribución marginal estacionaria del proceso GARCH(1,1).

La existencia de una única solución posible a la ecuación (2.10) para un  $k \geq 0$ , puede comprobarse teniendo en cuenta que  $\psi(k, \boldsymbol{\theta} | \eta)$  es una función continua y estrictamente convexa en  $k$ ,  $\psi(0, \boldsymbol{\theta} | \eta) = 0$  y  $\psi(k, \boldsymbol{\theta} | \eta) = \infty$  para  $k = k_0$  suficientemente grande, de forma que esta función admite una única solución para algún  $0 < \kappa < k_0$ .

Existen pocos casos donde esta ecuación toma una forma analítica explícita que permite encontrar numéricamente la solución de forma relativamente simple. En el caso ARCH(1),  $\beta_1 = 0$ , (de Haan *et.al.* (1989)) se tiene que la solución  $\kappa$  viene dada por la solución no nula de  $E[z_0^{2\kappa}] = \alpha_1^{-\kappa}$  que, en el caso de distribución normal para las innovaciones, queda de la forma  $\Gamma(\kappa + 1/2) = \sqrt{\pi}(2\alpha_1)^{-\kappa}$ .

En el caso GARCH(1,1) integrado (IGARCH(1,1)), la única solución para todos las combinaciones de valores tales que  $\alpha_1 + \beta_1 = 1$  es para  $\kappa = 1$ ,

$$E[(\alpha_1 z_0^2 + \beta_1)^{\kappa}] = E[(1 + \alpha_1(z_0^2 - 1))^{\kappa}] = E[(z_0^2 + \beta_1(1 - z_0^2))^{\kappa}] = 0 \Leftrightarrow \kappa = 1.$$

En cualquier otro caso, no existe una solución analítica para esta ecuación, de forma que debe ser evaluada numéricamente para determinar el valor (aproximado) de  $\kappa$ .

## 2.4 Condición de existencia de momentos de la distribución marginal

Bajo la condición de estacionariedad estricta, los momentos de orden  $2m$  de la distribución marginal de  $\varepsilon_t$ , para  $m \in (0, \infty)$ , vienen dados por

$$E(\varepsilon_t^{2m}) = E(z_t^{2m})E(h_t^m(\boldsymbol{\theta})) \tag{2.11}$$

donde, a partir de (2.5) y en el caso de innovaciones *i.i.d.* con momento  $2m$ -ésimo

finito, se tiene que por aplicación repetida de la desigualdad de Minkowski que

$$(E(h_t^m(\boldsymbol{\theta})))^{1/m} = \alpha_0 \left( E \left( \left\{ \sum_{i=1}^{\infty} \prod_{j=1}^{i-1} A_{t-j}(\boldsymbol{\theta}) \right\}^m \right) \right)^{1/m} \leq \alpha_0 \sum_{i=0}^{\infty} (E(A_0^m(\boldsymbol{\theta})))^{i/m}$$

existe sólo si la serie de potencias final es convergente, es decir, si

$$\gamma_m(\boldsymbol{\theta} | \eta) = E(A_0^m(\boldsymbol{\theta})) = E((\alpha_1 z_0^2 + \beta_1)^m) < 1 \quad (2.12)$$

para todo  $m \in (1, \infty)^1$ . Por tanto, (2.10) determina el máximo valor de  $m$  para el que se verifica la condición (2.12) de existencia de momento. Por el contrario, la desigualdad

$$E((\alpha_1 z_0^2 + \beta_1)^k) \geq 1$$

implica que tanto  $E(h_t^k(\boldsymbol{\theta}))$  como  $E(\varepsilon_t^{2k})$  son infinitos, puesto que a partir de la representación como serie binomial de  $E(h_t^k(\boldsymbol{\theta}))$  se tiene que

$$\begin{aligned} E(h_t^k(\boldsymbol{\theta})) &= E[(\alpha_0 + A_{t-1}(\boldsymbol{\theta})h_{t-1}(\boldsymbol{\theta}))^k] = \sum_{k=0}^{\infty} \binom{k}{k} \alpha_0^k E[A_{t-1}^{k-k}(\boldsymbol{\theta})h_{t-1}^{k-k}(\boldsymbol{\theta})] \\ &= E[A_{t-1}^k(\boldsymbol{\theta})]E[h_{t-1}^k(\boldsymbol{\theta})] + \sum_{k=1}^{\infty} \binom{k}{k} \alpha_0^k E[A_{t-1}^{k-k}(\boldsymbol{\theta})]E[h_{t-1}^{k-k}(\boldsymbol{\theta})] \end{aligned}$$

y dado que  $\alpha_0 > 0$ ,

$$E(h_t^k(\boldsymbol{\theta})) > E[A_{t-1}^k(\boldsymbol{\theta})]E[h_{t-1}^k(\boldsymbol{\theta})] = E[h_t^k(\boldsymbol{\theta})]E[A_t^k(\boldsymbol{\theta})]$$

lo que implica que esta relación sólo es posible si  $E(h_t^k(\boldsymbol{\theta})) = \infty$ . Si  $m$  es un entero positivo, entonces (2.12) implica que para todo  $k \leq m$   $\gamma_k(\boldsymbol{\theta} | \eta) < 1$ , de forma que

$$E(h_t^m(\boldsymbol{\theta})) = \frac{1}{1 - \gamma_m(\boldsymbol{\theta} | \eta)} \sum_{k=1}^m \binom{m}{k} \alpha_0^k E(h_t^{m-k}(\boldsymbol{\theta})) \gamma_{m-k}(\boldsymbol{\theta} | \eta) < \infty$$

(ver Bollerslev (1986), He y Teräsvirta (1999), Ling y McAleer (2002) y Liu (2006)). Nelson (1990) (Teorema 6) proporciona expresiones explícitas para el momento de orden  $m$  de  $A_0(\boldsymbol{\theta})$  en el proceso GARCH(1,1) en el caso de distribución normal y Cauchy estándar. Una consecuencia inmediata de este resultado es que un proceso GARCH no tiene momentos finitos de todos los órdenes, incluso si la distribución de las innovaciones  $z_0$  tiene soporte no acotado (es decir, momentos finitos de todos los órdenes), de forma que en general los procesos GARCH tendrán colas pesadas. De forma similar al resultado anterior, el momento  $m$ -ésimo de la distribución marginal

<sup>1</sup> Además, por la desigualdad de Jensen para funciones convexas, se tiene que

$$E(A_0^m(\boldsymbol{\theta})) = E(A_0(\boldsymbol{\theta})A_0^{m-1}(\boldsymbol{\theta})) \geq E(A_0(\boldsymbol{\theta}))E(A_0^{m-1}(\boldsymbol{\theta}))$$

Así,  $E(A_0^m(\boldsymbol{\theta})) < 1$  implica que  $E(A_0^{m-1}(\boldsymbol{\theta})) < 1$ .

estacionaria de  $\varepsilon_t$  viene dado por  $E(|\varepsilon_t|^m) = E(|z_t|^m)E(h_t^{m/2}(\boldsymbol{\theta}))$ , que será finito si y sólo si  $\gamma_{k/2}(\boldsymbol{\theta}, \eta) = E[A_t^{k/2}(\boldsymbol{\theta})] < 1, 1 \leq k \leq m$ .

Por definición de un proceso GARCH, la distribución de las innovaciones  $z_t$  es desconocida, de forma que supuestos sobre las características de la misma son puramente hipotéticas. Sin embargo, éstas determinan propiedades fundamentales de la distribución asintótica de los estimadores de los parámetros GARCH y de estadísticos tales como las funciones de autocovarianza y autocorrelación muestral del cuadrado de un proceso GARCH, como han establecido Mikosch, T. y C. Stàricà (2000) (ver también los trabajos recientes de Hall y Yao (2003) y Mikosch y Straumann (2006) sobre los resultados aplicables en los casos de innovaciones con distribuciones de colas pesadas y estables o de varianza infinita, respectivamente). Berkes y Horváth (2003) establecen la consistencia y normalidad asintótica del estimador QMV bajo normalidad condicional de los parámetros GARCH bajo la condición de existencia de, al menos, la varianza ( $\kappa > 1$ ) y momento de cuarto orden ( $\kappa > 2$ ), respectivamente, de la distribución de las innovaciones y establecen la conexión entre la tasa de consistencia de este estimador y una condición de existencia de momento fraccional de  $z_t$  de la forma,

$$T^{1-1/\nu} |\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}| \xrightarrow{c.s.} 0 \Leftrightarrow E |z_0^2|^\nu < \infty, 1 < \nu < 2$$

que en la práctica se verificará a partir de la condición de existencia de momento sólo si  $\kappa \geq 2$ . Por otro lado, Berkes y Horváth (2004) establecen resultados similares en el caso de la estimación QMV bajo distribuciones condicionales distintas de la normal. Bajo un conjunto de condiciones de regularidad verificables sobre las características de la distribución asumida para las innovaciones y la existencia de momentos de orden superior al segundo, establecen la consistencia casi segura del estimador QMV de  $\boldsymbol{\theta}$ . Por todas estas razones, resulta muy importante en la práctica verificar la condición de estacionariedad estricta (directamente o bien mediante la condición más fuerte de estacionariedad en covarianza) y tratar de determinar el valor del parámetro  $\kappa$  de índice de cola o de grosor de cola de la distribución estacionaria.

### **3. Estimación del parámetro de grosor de cola**

Existen diversos estudios, teóricos y empíricos, relativos a la estimación del denominado parámetro de índice de cola que determina el comportamiento básico de la función de cola superior de una distribución regularmente variable, que es una

condición necesaria y suficiente para que dicha distribución pertenezca al dominio máximo de atracción de una distribución de valor extremo tipo Fréchet,  $H(\alpha)$ ,  $\alpha > 0$  de la que forman parte las distribuciones de colas anchas, de gran interés en aplicaciones financieras. La gran mayoría de estos trabajos emplean la aproximación semiparamétrica de Hill (1975), basada en el supuesto de que la distribución *i.i.d.* subyacente de los valores extremos cae en el dominio máximo de atracción de la distribución de valor extremo de Fréchet. La estimación del índice de cola basada en máxima verosimilitud condicional implica la elección de un umbral que determina la fracción de la muestra de observaciones extremas requeridas para implementar el estimador y que resulta ser esencial para el trade-off entre sesgo y varianza. Huisman *et.al.* (2001) proponen un promedio ponderado de estimadores de Hill para diferentes valores del umbral que permite corregir por el sesgo de infraestimación en muestras finitas, mientras que Wagner y Marsh (2005) examinan las propiedades en muestras finitas de estimadores tipo Hill adaptativos bajo la clase de distribuciones marginales *t*-student, incluyendo modelos GARCH, y proponen una aproximación para la corrección por sesgo de las estimaciones basadas en el modelo considerado. En el contexto general de la aplicación de este estimador a series temporales financieras, que generalmente presentan dependencia de largo plazo, es que la teoría asintótica desarrollada se basa en el supuesto *i.i.d.* de forma que, como obtienen Kearns y Pagan (1997) entre otros, además del sesgo en muestras finitas, la varianza de la distribución asintótica convencional sobreestima la varianza del estimador en este caso. Hsing (1991) establece que bajo dependencia serial lineal, el estimador de Hill es aún consistente y asintóticamente normal con las mismas tasas de convergencia que en el caso *i.i.d.*, pero con diferente varianza. Quintos *et.al.* (2001) demuestran que en el caso de dependencia ARCH(1) o GARCH(1,1), el estimador de Hill tiene la misma distribución asintótica que en el caso de dependencia lineal. Sin embargo, en el contexto de procesos tipo GARCH, caracterizados por una distribución marginal estacionaria de colas anchas, la evidencia en estos y en otros muchos trabajos indica diferencias sustanciales en las estimaciones resultantes de este parámetro y la sensibilidad de los resultados obtenidos a la estructura del modelo GARCH en términos de los valores de los parámetros y distribución asumida para las innovaciones. La posible inconsistencia de estos resultados puede ser debida a que el estimador de Hill (y sus variantes) capturan únicamente el comportamiento de las colas de la distribución condicional, y esto no parece ser suficiente para una buena descripción de la verosimilitud de valores

extremos, de forma que parece más apropiado emplear la distribución condicional en su conjunto para describir las colas de la distribución incondicional y de aquí la de los datos observados. Esta es la base de la propuesta inicial de Starica y Pictet (1997), luego refinada por Mikosch y Starica (2000), Berkes, Horváth y Kokoszka (2003) e Iglesias y Linton (2007) que describimos a continuación.

### 3.1 El estimador de Starica y Pictet (1997)

Starica y Pictet (1997) proponen obtener  $\kappa$  evaluando directamente la condición (2.10), mediante integración numérica, asumiendo una distribución particular para las innovaciones y empleando en lugar de  $\alpha_1, \beta_1$  las estimaciones de estos coeficientes por máxima verosimilitud bajo normalidad condicional, es decir,

$$\psi(\kappa, \hat{\boldsymbol{\theta}}_T | \eta) = E[(\hat{\alpha}_{1,T} z_0^2 + \hat{\beta}_{1,T})^\kappa] - 1 = 0 \quad (3.1)$$

de forma que la solución  $\hat{\kappa}_T = \kappa(\hat{\boldsymbol{\theta}}_T | \eta)$  depende sólo parcialmente de los datos a través de las estimaciones de  $\boldsymbol{\theta}$ . Tomando éstas como dadas, se trata entonces de encontrar la solución de<sup>2</sup>

$$E[(\hat{\alpha}_{1,T} z_0^2 + \hat{\beta}_{1,T})^\kappa] = \sum_{k=0}^{\infty} \binom{\kappa}{k} \hat{\alpha}_{1,T}^k \hat{\beta}_{1,T}^{\kappa-k} E[z_0^{2k}] = 1$$

en términos de  $\kappa$ . Puesto que la función (3.1) tiene derivadas continuas de primer orden en el entorno de  $(\hat{\alpha}_{1,T}, \hat{\beta}_{1,T})$  y derivada de primer orden no nula en  $(\hat{\alpha}_{1,T}, \hat{\beta}_{1,T})$ , por aplicación del teorema de la Función Implícita se puede concluir que esta solución es única y permite determinar la distribución asintótica normal del estimador resultante, aunque con una tasa de convergencia no especificada. Esta propuesta presenta una serie de inconvenientes que comentamos a continuación.

A. Se requiere la evaluación numérica aproximada de la integral resultante para un conjunto de valores admisibles de  $\kappa$ ,  $\kappa \in (0, \kappa_{\max})$ , que, en general puede requerir un cierto esfuerzo computacional, especialmente para ciertas distribuciones de las innovaciones.

B. En el mismo sentido, se requiere imponer un supuesto distribucional para las innovaciones para evaluar el valor esperado, de forma que podría hablarse de pseudo o

---

<sup>2</sup> Puesto que los coeficientes binomiales  $\binom{\kappa}{k}$  se definen para un  $\kappa$  no entero en general, los factoriales se expresan en términos de la función gamma.

quasi-estimación numérica de  $\kappa$ , que puede diferir del resultado obtenido bajo la “verdadera” distribución de las innovaciones.

C. Para tener en cuenta la información contenida en los datos respecto del parámetro  $\kappa$  en el caso de parámetros GARCH estimados es preciso no sólo emplear sus estimaciones en lugar de  $\alpha_1, \beta_1$  en (3.1) sino que también se requiere emplear las innovaciones estimadas,  $z_t(\hat{\boldsymbol{\theta}}_T) = \varepsilon_t / \sqrt{h_t(\hat{\boldsymbol{\theta}}_T)}$ , en lugar de las hipotéticas  $z_t = z_t(\boldsymbol{\theta})$ , de forma que la solución se obtendría mediante evaluación numérica de

$$\tilde{\Psi}(\kappa, \hat{\boldsymbol{\theta}}_T | \eta) = E[(\hat{\alpha}_{1,T} z_t^2(\hat{\boldsymbol{\theta}}_T) + \hat{\beta}_{1,T})^\kappa] - 1 = 0$$

empleando integración numérica aproximada o bien alguna estimación no paramétrica de la densidad de las innovaciones. De igual forma, cambiaría la distribución asintótica del estimador, puesto que a partir de la forma linealizada

$$\kappa(\hat{\boldsymbol{\theta}}_T | \eta) = \kappa(\boldsymbol{\theta}_0 | \eta) + (\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0)' \left. \frac{\partial \kappa(\hat{\boldsymbol{\theta}}_T | \eta)}{\partial \hat{\boldsymbol{\theta}}_T} \right|_{\hat{\boldsymbol{\theta}}_T = \boldsymbol{\theta}_0}$$

se tiene, por el teorema de la Función Implícita, que

$$\frac{\partial \kappa(\hat{\boldsymbol{\theta}}_T | \eta)}{\partial \hat{\boldsymbol{\theta}}_T} = - \left( \frac{\partial \tilde{\Psi}(\kappa, \hat{\boldsymbol{\theta}}_T | \eta)}{\partial \kappa} \right)^{-1} \frac{\partial \tilde{\Psi}(\kappa, \hat{\boldsymbol{\theta}}_T | \eta)}{\partial \hat{\boldsymbol{\theta}}_T}$$

y

$$\frac{\partial \tilde{\Psi}(\kappa, \hat{\boldsymbol{\theta}}_T | \eta)}{\partial \hat{\boldsymbol{\theta}}_T} = \kappa E \left( (\hat{\alpha}_{1,T} z_t^2(\hat{\boldsymbol{\theta}}_T) + \hat{\beta}_{1,T})^{\kappa-1} \frac{\partial A_t(\hat{\boldsymbol{\theta}}_T)}{\partial \hat{\boldsymbol{\theta}}_T} \right), \quad \frac{\partial A_t(\hat{\boldsymbol{\theta}}_T)}{\partial \hat{\boldsymbol{\theta}}_T} = \begin{pmatrix} 0 \\ z_t^2(\hat{\boldsymbol{\theta}}_T) \\ 1 \end{pmatrix} + \hat{\alpha}_{1,T} \frac{\partial z_t^2(\hat{\boldsymbol{\theta}}_T)}{\partial \hat{\boldsymbol{\theta}}_T}$$

admitiendo siempre la posibilidad de intercambiar los signos integral y valor esperado. Por todo esto, esta propuesta interesante presenta bastante inconvenientes que hacen cuestionable su utilización práctica, salvo como primera aproximación.

### 3.2 El estimador de Iglesias y Linton (1997)

Mikosch y Starica (2000) realizan la primera propuesta de estimación de  $\kappa$  a través del análogo empírico de (3.1), definido como

$$\Psi_T(\kappa, \hat{\boldsymbol{\theta}}_T) = \frac{1}{T} \sum_{t=1}^T A_t^\kappa(\hat{\boldsymbol{\theta}}_T) - 1, \quad A_t(\hat{\boldsymbol{\theta}}_T) = \hat{\alpha}_{1,T} z_t^2(\hat{\boldsymbol{\theta}}_T) + \hat{\beta}_{1,T} \quad (3.2)$$

Berkes, Horváth y Kokoszka (2003) siguen la propuesta de Mikosch y Starica (2000) y refinan los resultados de consistencia y distribución asintótica normal del estimador de  $\kappa$  utilizando el teorema ergódico y el teorema central del límite para martingalas.

Posteriormente, Iglesias y Linton (2007) (IL), generalizan esta propuesta al caso de la estimación del parámetro de grosor de cola  $\kappa$  en procesos ARMA-GARCH y GARCH asimétricos que satisfacen la misma estructura de representación SRE<sup>3</sup> que el proceso GARCH(1,1) en la sección 2.1. Así, el estimador  $\hat{\kappa}_T = \kappa(\hat{\boldsymbol{\theta}}_T)$  se define como la única solución de la ecuación estimable

$$\Psi_T(\kappa, \hat{\boldsymbol{\theta}}_T) = \frac{1}{T} \sum_{t=1}^T A_t^{\kappa/2}(\hat{\boldsymbol{\theta}}_T) - 1 = \frac{1}{T} \sum_{t=1}^T (\hat{\alpha}_{1,T} z_t^2(\hat{\boldsymbol{\theta}}_T) + \hat{\beta}_{1,T})^{\kappa/2} - 1 \quad (3.3)$$

bajo la condición numérica de cálculo,

$$\hat{\kappa}_T = \kappa(\hat{\boldsymbol{\theta}}_T) : \Psi_T(\hat{\kappa}_T, \hat{\boldsymbol{\theta}}_T) = o_p(T^{-1/2}). \quad (3.4)$$

En este caso, el parámetro  $\kappa$ , como solución de (3.3)-(3.4), es el parámetro de grosor de cola de la distribución estacionaria de  $(\varepsilon_t, \sqrt{h_t(\boldsymbol{\theta})})$ , que se corresponde con un parámetro  $(\kappa/2)$  de la distribución estacionaria regularmente variable de  $(\varepsilon_t^2, h_t(\boldsymbol{\theta}))$ . La solución de (3.3)-(3.4) puede obtenerse mediante un simple proceso de búsqueda sobre una rejilla de valores apropiados para  $\kappa$ ,  $\mathbf{K} \in \mathbb{R}_{(0)}^+$ .

Bajo un conjunto de condiciones de regularidad estándar sobre la distribución de las innovaciones y el comportamiento asintótico del estimador de  $\boldsymbol{\theta}$  que se satisfacen tanto en el caso de innovaciones *i.i.d.* como en el caso del modelo GARCH semi-fuerte en relación con la estimación QMV, Iglesias y Linton (2007) establecen la normalidad asintótica del estimador propuesto y la tasa de convergencia,  $T^{1/2}$ , es más rápida que la de los estimadores tipo Hill, de forma que es asintóticamente más eficiente. Además, desde el punto de vista tanto teórico como empírico, resulta recomendable emplear en (3.2) o (3.3) las innovaciones estimadas reescaladas puesto que forma de la ecuación estimable está sujeta a dos restricciones básicas de la distribución marginal de  $z_t$ :  $E(z_t) = 0$  y  $Var(z_t) = 1$ .

En el siguiente epígrafe estudiaremos algunas características del comportamiento del estimador de  $\kappa$  de IL (2007) que pueden servir de utilidad especialmente al analizar los resultados de un estudio empírico con series financieras reales bajo el supuesto de comportamiento GARCH. En el epígrafe 5 ilustraremos los resultados de este análisis mediante un experimento de simulación y, por último, en el epígrafe 6 presentaremos los resultados de la aplicación a tres series de rendimientos diarios de tipos de cambio.

---

<sup>3</sup> En particular, consideran su aplicabilidad al caso del modelo GARCH asimétrico GJR-GARCH(1,1) de Glosten, Jagannathan y Runkle (1993).

## 4. Comportamiento del estimador IL (2007) de $\kappa$

### 4.1 Parámetros GARCH conocidos vs estimados

En primer lugar, analizamos el efecto de utilizar los parámetros GARCH estimados y, especialmente, las innovaciones estimadas en comparación con el caso hipotético de parámetros e innovaciones conocidas para explicar los resultados obtenidos en las simulaciones.

Si los parámetros GARCH fuesen conocidos así como las innovaciones, el estimador de  $\kappa$  se obtendría como solución de la ecuación

$$\Psi_T(\kappa, \boldsymbol{\theta}) = \frac{1}{T} \sum_{t=1}^T A_t^\kappa(\boldsymbol{\theta}) = \beta_1^\kappa \sum_{k=0}^{\infty} \binom{\kappa}{k} (\alpha_1 / \beta_1)^k \left( \frac{1}{T} \sum_{t=1}^T z_t^{2k}(\boldsymbol{\theta}) \right) = \beta_1^\kappa \sum_{k=0}^{\infty} \binom{\kappa}{k} (\alpha_1 / \beta_1)^k \mu_{2k,T}(\boldsymbol{\theta}) - 1$$

que depende tanto de la combinación de  $(\alpha_1, \beta_1)$  como de los momentos muestrales de orden  $2k$  de las innovaciones teóricas. En el caso de parámetros GARCH estimados, las innovaciones estimadas pueden expresarse en función de las teóricas a través de la relación  $z_t(\hat{\boldsymbol{\theta}}_T) = z_t(\boldsymbol{\theta}) \sqrt{h_t(\boldsymbol{\theta}) / h_t(\hat{\boldsymbol{\theta}}_T)}$ , de forma que un desarrollo en serie de Taylor de primer orden en torno a  $\boldsymbol{\theta}$  permite obtener la expresión

$$z_t(\hat{\boldsymbol{\theta}}_T) = z_t(\boldsymbol{\theta}) - \frac{1}{2} z_t(\boldsymbol{\theta}) \frac{\partial \ln h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} (\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) = z_t(\boldsymbol{\theta}) - \frac{1}{2} z_t(\boldsymbol{\theta}) \hat{\phi}_{t,T}(\boldsymbol{\theta}) \quad (4.1)$$

de forma que la potencia  $m$ -ésima de (4.1) viene dada por

$$z_t^m(\hat{\boldsymbol{\theta}}_T) = \frac{z_t^m(\boldsymbol{\theta})}{2^m} \sum_{k=0}^m 2^k (-1)^{m-k} \binom{m}{k} \hat{\phi}_{t,T}^{m-k}(\boldsymbol{\theta})$$

Así, el momento muestral de orden  $2k$  de  $z_t(\hat{\boldsymbol{\theta}}_T)$  puede expresarse como

$$\mu_{2k,T}(\hat{\boldsymbol{\theta}}_T) = \frac{1}{T} \sum_{t=1}^T z_t^{2k}(\hat{\boldsymbol{\theta}}_T) = \frac{1}{2^{2k}} \sum_{m=0}^{2k} 2^m (-1)^{2k-m} \binom{2k}{m} \left( \frac{1}{T} \sum_{t=1}^T z_t^{2k-m}(\boldsymbol{\theta}) \hat{\phi}_{t,T}^{2k-m}(\boldsymbol{\theta}) \right)$$

es decir,

$$\mu_{2k,T}(\hat{\boldsymbol{\theta}}_T) = \frac{1}{2^{2k}} \left\{ \mu_{2k,T}(\boldsymbol{\theta}) + \frac{1}{T} \sum_{t=1}^T z_t^{2k}(\boldsymbol{\theta}) (\hat{\phi}_{t,T}^{2k}(\boldsymbol{\theta}) - 1) + \sum_{m=1}^{2k} 2^m (-1)^{2k-m} \binom{2k}{m} \left( \frac{1}{T} \sum_{t=1}^T z_t^{2k-m}(\boldsymbol{\theta}) \hat{\phi}_{t,T}^{2k-m}(\boldsymbol{\theta}) \right) \right\}$$

donde el factor  $1/2^{2k}$  puede provocar que estos momentos puedan hacerse relativamente pequeños rápidamente de forma que la contribución de términos de orden  $k$  elevado es pequeña y, por tanto, la ecuación

$$\Psi_T(\kappa, \hat{\boldsymbol{\theta}}_T) = \frac{1}{T} \sum_{t=1}^T A_t^\kappa(\hat{\boldsymbol{\theta}}_T) = \hat{\beta}_{1,T}^\kappa \sum_{k=0}^{\infty} \binom{\kappa}{k} (\hat{\alpha}_{1,T} / \hat{\beta}_{1,T})^k \mu_{2k,T}(\hat{\boldsymbol{\theta}}_T) - 1 \quad (4.2)$$

se aproxima a cero para un valor de  $\kappa$  más pequeño, es decir,  $\kappa(\hat{\boldsymbol{\theta}}_T) < \kappa(\boldsymbol{\theta})$ . Así, la estimación de los parámetros GARCH determina una infraestimación de la magnitud del parámetro de grosor de cola  $\kappa$ . Se puede obtener la misma conclusión empleando la siguiente representación de la ecuación estimable (4.2)

$$\Psi_T(\kappa, \hat{\boldsymbol{\theta}}_T) = \Psi_T(\kappa, \boldsymbol{\theta}) + (\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})' \frac{1}{T} \sum_{t=1}^T \frac{\partial A_t^\kappa(\hat{\boldsymbol{\theta}}_T)}{\partial \hat{\boldsymbol{\theta}}_T} \Big|_{\hat{\boldsymbol{\theta}}_T = \boldsymbol{\theta}} \quad (4.3)$$

donde

$$\frac{\partial A_t^\kappa(\hat{\boldsymbol{\theta}}_T)}{\partial \hat{\boldsymbol{\theta}}_T} = \kappa A_t^{\kappa-1}(\hat{\boldsymbol{\theta}}_T) \left\{ \begin{pmatrix} 0 \\ z_t^2(\hat{\boldsymbol{\theta}}_T) \\ 1 \end{pmatrix} + \hat{\alpha}_{1,t} \frac{\partial z_t^2(\hat{\boldsymbol{\theta}}_T)}{\partial \hat{\boldsymbol{\theta}}_T} \right\}$$

y

$$\frac{\partial z_t^2(\hat{\boldsymbol{\theta}}_T)}{\partial \hat{\boldsymbol{\theta}}_T} = -z_t^2(\hat{\boldsymbol{\theta}}_T) \frac{\partial \ln h_t(\hat{\boldsymbol{\theta}}_T)}{\partial \hat{\boldsymbol{\theta}}_T}$$

Así, se obtiene que el término final del segundo miembro de (4.3) tiene la forma,

$$(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})' \frac{1}{T} \sum_{t=1}^T \frac{\partial A_t^\kappa(\hat{\boldsymbol{\theta}}_T)}{\partial \hat{\boldsymbol{\theta}}_T} \Big|_{\hat{\boldsymbol{\theta}}_T = \boldsymbol{\theta}} = \frac{\kappa}{T} \sum_{t=1}^T A_t^{\kappa-1}(\boldsymbol{\theta}) \left\{ (\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})' \begin{pmatrix} 0 \\ z_t^2(\boldsymbol{\theta}) \\ 1 \end{pmatrix} - \alpha_1 z_t^2(\boldsymbol{\theta}) \phi_{t,T}(\boldsymbol{\theta}) \right\}$$

que, en general, tiene siempre valor esperado no nulo, de forma que para el valor de  $\kappa$  que determina que  $E(\Psi_T(\kappa, \boldsymbol{\theta})) = 0$ , en (4.2) se tiene que  $E(\Psi_T(\kappa, \hat{\boldsymbol{\theta}}_T)) > 0$ , aunque cabe esperar que para un tamaño muestral incluso moderado predomine la consistencia de la estimación de  $\boldsymbol{\theta}$  y ambas soluciones estén muy próximas.

## 4.2 GARCH(1,1) vs GARCH(p,q)

En segundo lugar, analizamos el efecto de un error de especificación en la ecuación de volatilidad condicional GARCH. Así, si se supone que el proceso generador de  $\varepsilon_t$  es GARCH(1,1), se tiene que las innovaciones vienen dadas por

$$z_t^2(\boldsymbol{\theta}_1) = \frac{\varepsilon_t^2}{h_t(\boldsymbol{\theta}_1)} = z_t^2(\boldsymbol{\theta}_0) \frac{h_t(\boldsymbol{\theta}_0)}{h_t(\boldsymbol{\theta}_1)} = z_t^2(\boldsymbol{\theta}_0) \cdot R_t(\boldsymbol{\theta}), \quad \boldsymbol{\theta}_1 = (a_0, a_1, b_1) \quad (4.4)$$

con

$$A_t(\boldsymbol{\theta}_1) = a_1 z_t^2(\boldsymbol{\theta}_1) + b_1 = a_1 z_t^2(\boldsymbol{\theta}_0) \cdot R_t(\boldsymbol{\theta}) + b_1, \quad (4.5)$$

mientras que si la correcta especificación de la varianza condicional es GARCH(p,q) ( $p, q \geq 1$ ),

$$h_t(\boldsymbol{\theta}_0) = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2(\boldsymbol{\theta}_0) + \sum_{j=1}^q \beta_j h_{t-j}(\boldsymbol{\theta}_0) = \alpha_0 + \sum_{i=1}^m A_{i,t-i}(\boldsymbol{\theta}_0) h_{t-i}(\boldsymbol{\theta}_0)$$

con  $\varepsilon_{t-i}^2(\boldsymbol{\theta}_0) = z_{t-i}^2(\boldsymbol{\theta}_0) h_{t-i}(\boldsymbol{\theta}_0)$  y  $A_{i,t-i}(\boldsymbol{\theta}_0) = \alpha_i z_{t-i}^2(\boldsymbol{\theta}_0) + \beta_j$ , es decir,

$$h_t(\boldsymbol{\theta}_0) = \alpha_0 + A_{1,t-1}(\boldsymbol{\theta}_0) h_{t-1}(\boldsymbol{\theta}_0) + \sum_{i=2}^m A_{i,t-i}(\boldsymbol{\theta}_0) h_{t-i}(\boldsymbol{\theta}_0)$$

donde  $z_t(\boldsymbol{\theta}_0)$ , en el caso GARCH fuerte, representa las innovaciones con media nula y varianza unitaria. Entonces, a partir de (4.5), se tiene que

$$E[A_t^\kappa(\boldsymbol{\theta}_1)] = b_1^\kappa \sum_{k=0}^{\infty} \binom{\kappa}{k} (a_1/b_1)^k \mu_{2k}(\boldsymbol{\theta}_0) \cdot E[R_t^\kappa(\boldsymbol{\theta})] \quad (4.6)$$

será en general distinto del valor esperado bajo el supuesto de correcta especificación de la varianza condicional, puesto que en ese caso  $R_t(\boldsymbol{\theta}) = 1$ . Considerando la siguiente relación entre los parámetros GARCH(1,1) y GARCH( $p,q$ ) para la constante y los término del primer retardo,  $a_0 = \alpha_0 + \delta_0$ ,  $a_1 = \alpha_1 + \delta_1$ ,  $b_1 = \beta_1 + \delta_2$ , y llamando  $\xi_t(\boldsymbol{\theta}) = h_t(\boldsymbol{\theta}_0) - h_t(\boldsymbol{\theta}_1)$ , se tiene que

$$\xi_t(\boldsymbol{\theta}) = \beta_1 \xi_{t-1}(\boldsymbol{\theta}) - h_t(\boldsymbol{\delta}_1) + \sum_{i=2}^m A_{i,t-i}(\boldsymbol{\theta}_0) h_{t-i}(\boldsymbol{\theta}_0) = \beta_1 \xi_{t-1}(\boldsymbol{\theta}) - h_t(\boldsymbol{\delta}_1) + g_{t,m}(\boldsymbol{\theta}_0)$$

donde  $h_t(\boldsymbol{\delta}_1) = \delta_0 + \delta_1 \varepsilon_{t-1}^2(\boldsymbol{\theta}_0) + \delta_2 h_{t-1}(\boldsymbol{\theta}_1)$ . Entonces, mediante sustituciones recursivas se tiene que la diferencia entre la correcta varianza condicional GARCH( $p,q$ ) y la especificada GARCH(1,1) viene dada por

$$\xi_t(\boldsymbol{\theta}) = \beta_1^t \xi_0(\boldsymbol{\theta}) + \sum_{i=1}^t \beta_1^{i-1} (g_{t-i+1,m}(\boldsymbol{\theta}_0) - h_{t-i+1}(\boldsymbol{\delta}_1)) \quad (4.7)$$

que será siempre positiva si los parámetro de GARCH(1,1) toman los mismos valores que los correspondientes en GARCH( $p,q$ ) (es decir, si  $\delta_i = 0$ ,  $i = 0,1,2$ ), de forma que el ratio  $R_t(\boldsymbol{\theta}) > 1 \forall t$  y, por tanto, el error de especificación supondrá una sobreestimación de  $\kappa$ . Sin embargo, la evidencia apunta a que con un error de especificación de esta naturaleza, los coeficientes  $\delta_i > 0$ , puesto que con el modelo GARCH(1,1) especificado incorrectamente se recogerá parte de la dinámica omitida a través de un incremento del valor de los parámetros. En este caso, no se puede apuntar un efecto general.

### 4.3 Modelo GARCH(1,1) semi-fuerte

En este caso, las innovaciones del proceso GARCH, definidas como los términos de error re-escalados, son estrictamente estacionarias y ergódicas, pero no necesariamente

independientes. Así, si bien la condición de estacionariedad estricta sigue siendo la misma que en el caso GARCH fuerte, la condición de existencia de momentos de la distribución marginal y la estructura de momentos no coincide porque en este caso los términos  $A_{t-j}(\boldsymbol{\theta}) = \alpha_1 z_{t-j}^2(\boldsymbol{\theta}) + \beta_1$  no son ya independientes. Esto puede afectar a la magnitud de la estimación de  $\kappa$ , aunque la evidencia de simulación que se presenta posteriormente indica que el efecto se corrige al aumentar el tamaño muestral y en general supone una infraestimación de  $\kappa$  cuando se asume una estructura de dependencia débil para las innovaciones (AR(1) o MA(1) con coeficiente pequeño).

#### 4.4 Invarianza ante agregación temporal: leyes de escalamiento en procesos GARCH

Una forma de verificar empíricamente la naturaleza de la distribución estacionaria de variación regular de los procesos GARCH puede consistir en tratar de determinar la no invarianza de ciertas características de dicha distribución bajo agregación temporal, es decir, al cambiar la escala temporal de observación del proceso. Starica y Pictet (1997) examinan en este contexto el comportamiento de  $\kappa$  ante la escala temporal empleando el estimador de Hill (1975). En este epígrafe examinamos el comportamiento del estimador IL bajo agregación temporal del proceso de innovación de un proceso GARCH(1,1). En primer lugar, observamos que el momento muestral de orden  $2k$  de las innovaciones estimadas,  $\mu_{2k,T}(\hat{\boldsymbol{\theta}}_T)$ , puede escribirse de la forma

$$\mu_{2k,T}(\hat{\boldsymbol{\theta}}_T) = \sum_{t=1}^{\lfloor T/m \rfloor} \frac{1}{T} \sum_{j=1}^m z_{(t-1)m+j}^{2k}(\hat{\boldsymbol{\theta}}_T) = \sum_{t=1}^{\lfloor T/m \rfloor} \mu_{2k,t}(m, \hat{\boldsymbol{\theta}}_T)$$

de forma que en (4.2) se tiene que

$$\Psi_T(\kappa, \hat{\boldsymbol{\theta}}_T) = \hat{\beta}_{1,T}^\kappa \sum_{k=0}^{\infty} \binom{\kappa}{k} (\hat{\alpha}_{1,T} / \hat{\beta}_{1,T})^k \left\{ \sum_{t=1}^{\lfloor T/m \rfloor} \mu_{2k,t}(m, \hat{\boldsymbol{\theta}}_T) \right\} - 1$$

Puesto que (4.2) puede escribirse también como

$$\Psi_T(\kappa, \hat{\boldsymbol{\theta}}_T) = \frac{1}{T} \hat{S}_T^{(2k)}(1) \sum_{k=0}^{\infty} \binom{\kappa/2}{k} \hat{\alpha}_{1,T}^k \hat{\beta}_{1,T}^{\kappa/2-k} - 1 \quad (4.8)$$

donde

$$\hat{S}_T^{(2k)}(u) = \sum_{t=1}^{\lfloor Tu \rfloor} z_t^{2k}(\hat{\boldsymbol{\theta}}_T) \quad 0 \leq u \leq 1 \quad (4.9)$$

es el proceso de suma parcial del momento de orden  $2k$  basado en los residuos o

innovaciones estimadas en el modelo GARCH(1,1), definido y estudiado por Kulperber y Yu (2005) en el caso general GARCH( $p,q$ ). A partir de la representación (4.1), en el Teorema 1.1 de Kulperber y Yu (2005) se establece la convergencia uniforme del proceso de suma parcial de momento de orden  $k$  a un proceso Browniano más un término de corrección que dependen del momento del mismo orden de la distribución de las innovaciones del proceso GARCH y del término de error en la estimación de los parámetros GARCH de la forma

$$\frac{1}{\sqrt{T}} \hat{S}_T^{(k)}(u) \approx \frac{1}{\sqrt{T}} S_T^{(k)}(u) - \frac{k\mu_k}{2} \left( \frac{\partial \ln h_0(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \sqrt{T} (\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) \right) \quad (4.10)$$

donde  $S_T^{(k)}(u) = \sum_{t=1}^{\lfloor Tu \rfloor} z_t^k(\boldsymbol{\theta})$ ,  $0 \leq u \leq 1$  es el proceso de suma parcial de momento de las innovaciones del proceso GARCH que, bajo el supuesto i.i.d. converge a un proceso Browniano y,  $\mu_k = E[z_0^k]$ . Puesto que para  $u = 1$ , (4.9) puede escribirse como

$$\hat{S}_T^{(2k)}(1) = \sum_{t=1}^{\lfloor T/m \rfloor} [\hat{S}_T^{(2k)}(u_{t,T}(m)) - \hat{S}_T^{(2k)}(u_{t-1,T}(m))], \quad u_{t,T}(m) = t \left\lfloor \frac{m}{T} \right\rfloor \quad (4.11)$$

con  $\hat{S}_T^{(2k)}(u_{0,T}(m)) = \hat{S}_T^{(2k)}(0) = 0$ , y

$$\begin{aligned} \hat{S}_T^{(2k)}(u_{t,T}(m)) - \hat{S}_T^{(2k)}(u_{t-1,T}(m)) &= S_T^{(2k)}(u_{t,T}(m)) - S_T^{(2k)}(u_{t-1,T}(m)) \\ &\quad - k\mu_{2k} \lfloor m/T \rfloor \left( \sqrt{T} \frac{\partial \ln h_0(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \sqrt{T} (\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) \right) \end{aligned} \quad (4.12)$$

de forma que la diferencia entre los procesos de agregación de innovaciones no solapadas  $S_T^{(2k)}(u_{t,T}(m)) - S_T^{(2k)}(u_{t-1,T}(m))$  converge también a un proceso Browniano estándar (gaussiano), que presenta la propiedad de autoafinidad o autosimilaridad, de forma que bajo esta propiedad la distribución de este proceso depende básicamente de la escala de observación, es decir, del nivel de agregación.

Así, si se define la secuencia de innovaciones agregadas bajo el supuesto de que son variables flujo de la forma

$$z_{t(m)}(\boldsymbol{\theta}) = \sum_{j=1}^m z_{(t-1)m+j}(\boldsymbol{\theta}) \quad t = 1, \dots, \lfloor T/m \rfloor$$

entonces, bajo el principio de autoafinidad o autosimilaridad de la distribución de  $z_t$ , se tiene que

$$z_{t(m)}(\boldsymbol{\theta}) \stackrel{d}{=} m^\delta z_t(\boldsymbol{\theta})$$

donde  $\stackrel{d}{=}$  denota igualdad en distribución, y donde  $\delta = 1/\alpha$ , con  $\alpha = 2$  en el caso i.i.d. normal. Entonces si la secuencia de variables  $A_t(\boldsymbol{\theta})$  bajo agregación temporal de las

innovaciones, dada por  $A_{t(m)}(\boldsymbol{\theta}) = \alpha_1 z_{t(m)}^2(\boldsymbol{\theta}) + \beta_1$ , satisface la misma propiedad, entonces se puede escribir de forma aproximada

$$A_{t(m)}(\boldsymbol{\theta}) \stackrel{d}{=} \alpha_1 m^{2\delta} z_t^2(\boldsymbol{\theta}) + \beta_1 = A_t(\boldsymbol{\theta}) + \alpha_1 z_t^2(\boldsymbol{\theta})(m^{2\delta} - 1)$$

Si se denota por  $\kappa(\boldsymbol{\theta})$  la solución de la condición de existencia de momento máximo a partir del proceso  $A_t(\boldsymbol{\theta})$  (sin agregación) y por  $\kappa(m, \boldsymbol{\theta})$  a partir del proceso  $A_{t(m)}(\boldsymbol{\theta})$ , es decir, para el que se verifica que  $1 = E[A_{t(m)}^{\kappa(m, \boldsymbol{\theta})}(\boldsymbol{\theta})]$ , entonces se tiene que si se verifica la propiedad de autoafinidad,

$$E[A_{t(m)}^{\kappa(\boldsymbol{\theta})}(\boldsymbol{\theta})] > E[A_t^{\kappa(\boldsymbol{\theta})}(\boldsymbol{\theta})] = 1$$

de forma que debería observarse que  $\kappa(m, \boldsymbol{\theta}) < \kappa(\boldsymbol{\theta})$ , de forma que al aumentar el nivel de agregación disminuye el valor de  $\kappa$ . Sin embargo, esto contradice tanto los resultados teóricos presentados anteriormente como la evidencia empírica sobre la naturaleza de la distribución de un proceso GARCH al aumentar la frecuencia de observación.

El hecho de que el valor estimado de  $\kappa$ , tanto en el caso de parámetros GARCH conocidos como desconocidos, aumente con el nivel de agregación de las innovaciones (teóricas o estimadas) es totalmente consistente con el resultado de que la distribución marginal estrictamente estacionaria de los procesos GARCH es una distribución de variación regular, con colas gruesas, pero en general no estable, que viene dada parcialmente por la distribución de las innovaciones pero que tiene un carácter más complejo que el de éstas.

## 5. Resultados de simulación

En esta sección presentamos un conjunto de resultados de un experimento de simulación que trata de reflejar algunas de las características analizadas en la anterior sección. En todos los casos se ha simulado un modelo GARCH(1,1) con el mismo valor de los parámetros,  $(\alpha_0, \alpha_1, \beta_1) = (0.05, 0.15, 0.65)$ . Cada uno de los cuadros siguientes permite analizar varios aspectos analizados: el efecto de la estimación de los parámetros GARCH y de las innovaciones, el efecto de la distribución asumida para éstas (normal, t-student, distribución generalizada del error (GED) y mezcla de dos normales independientes). En todos los casos se han estimado los parámetros GARCH por MV bajo normalidad condicional (QMV). Se han considerado en general tamaños muestrales relativamente pequeños y se ha considerado la agregación de las

innovaciones bajo el supuesto tanto de variables flujo como stock (ver Drost y Nijman (1993)). En todos los casos se han realizado 500 replicaciones, y en todos los cuadrados aparece entre paréntesis el error estándar de simulación. Puesto que la estimación de  $\kappa$  se basa en la ecuación (3.3), se refiere al parámetro de grosor de cola de la distribución marginal tipo Pareto de  $\varepsilon_t$ . Los resultados pueden extenderse a la distribución de  $\varepsilon_t^2$  ó  $h_t(\boldsymbol{\theta})$  haciendo  $\kappa/2$ .

<b>Cuadro 1. Distribución N(0,1)</b>		Variable Flujo		Variable Stock	
		$\hat{\kappa}_T(m, \boldsymbol{\theta})$	$\hat{\kappa}_T(m, \hat{\boldsymbol{\theta}}_T)$	$\hat{\kappa}_T(m, \boldsymbol{\theta})$	$\hat{\kappa}_T(m, \hat{\boldsymbol{\theta}}_T)$
$T = 100$	$m = 1$	12.784 (3.053)	7.633 (6.499)		
	$m = 2$	13.998 (4.756)	7.832 (6.378)	13.757 (4.537)	8.053 (6.943)
$T = 250$	$m = 1$	11.655 (1.787)	9.715 (6.307)		
	$m = 2$	12.488 (2.674)	8.926 (6.601)	12.198 (2.518)	9.802 (6.843)
	$m = 5$	13.752 (4.635)	9.012 (6.463)	14.362 (4.787)	9.160 (6.926)
$T = 500$	$m = 1$	11.206 (1.369)	10.106 (5.425)		
	$m = 2$	11.756 (1.893)	10.437 (5.413)	11.617 (1.807)	10.534 (5.684)
	$m = 5$	12.747 (3.052)	10.569 (6.317)	12.781 (3.236)	11.214 (6.242)
	$m = 10$	13.889 (4.488)	10.694 (6.837)	13.903 (4.391)	11.230 (6.890)
$T = 1000$	$m = 1$	11.061 (0.982)	10.908 (4.295)		
	$m = 2$	11.182 (1.284)	11.153 (4.671)	11.216 (1.342)	10.919(3.961)
	$m = 5$	11.758 (2.106)	11.767 (4.751)	11.970 (2.103)	11.857 (4.541)
	$m = 10$	12.688 (2.950)	12.075 (5.278)	12.684 (2.692)	12.405 (5.391)
	$m = 20$	13.661 (4.409)	12.761 (6.491)	13.842 (4.655)	12.396 (6.699)

<b>Cuadro 2. Distribución T-Student, T(d)</b>		<b>T(d = 6)</b>		<b>T(d = 9)</b>	
Variable Flujo		$\hat{\kappa}_T(m, \boldsymbol{\theta})$	$\hat{\kappa}_T(m, \hat{\boldsymbol{\theta}}_T)$	$\hat{\kappa}_T(m, \boldsymbol{\theta})$	$\hat{\kappa}_T(m, \hat{\boldsymbol{\theta}}_T)$
$T = 100$	$m = 1$	8.585 (2.491)	5.911 (5.129)	9.755 (2.728)	6.732 (5.877)
	$m = 2$	11.644 (4.049)	7.048 (6.394)	12.550 (4.113)	6.873 (5.912)
$T = 250$	$m = 1$	7.409 (1.662)	7.213 (5.205)	8.524 (1.686)	7.707 (5.265)
	$m = 2$	9.249 (2.472)	7.488 (5.755)	10.309 (2.458)	8.451 (6.224)
	$m = 5$	12.803 (4.488)	8.282 (6.499)	13.317 (4.752)	9.114 (7.189)
$T = 500$	$m = 1$	6.814 (1.228)	6.967 (4.096)	8.107 (1.301)	8.166 (4.329)
	$m = 2$	8.475 (1.757)	8.434 (5.005)	9.635 (1.687)	9.202 (4.946)
	$m = 5$	10.923 (2.909)	9.108 (5.901)	11.705 (2.825)	10.155 (5.986)
	$m = 10$	13.117 (4.674)	9.742 (7.007)	13.386 (4.713)	10.691 (6.922)
$T = 1000$	$m = 1$	6.438 (1.007)	6.557 (2.757)	7.717 (1.016)	7.851 (2.955)
	$m = 2$	7.892 (1.413)	8.189 (3.559)	9.132 (1.264)	9.388 (4.259)
	$m = 5$	10.148(1.997)	9.897 (4.993)	10.836 (2.004)	10.373 (4.762)
	$m = 10$	11.585 (2.998)	10.343 (5.901)	12.334 (3.025)	11.301 (5.995)
	$m = 20$	13.612 (4.545)	11.309 (6.899)	13.574 (4.523)	11.451 (6.701)

En general, se observa el efecto de infraestimación de  $\kappa$  al emplear la estimación de los parámetros GARCH así como de las innovaciones  $z_t(\hat{\boldsymbol{\theta}}_T)$ , aunque incluso para tamaños muestrales moderados como  $T = 500$  o  $1000$  los resultados están muy próximos. Se observa que el sesgo es tanto mayor cuánto más leptocúrtica es la distribución de las

innovaciones, puesto que aún siendo consistentes las estimaciones de  $\theta$  basadas en QMV bajo normalidad condicional, existe una pérdida de eficiencia.

Por otro lado se observa el efecto directo sobre la magnitud de  $\kappa$  de la naturaleza de colas anchas o estrechas de las innovaciones. En relación con la agregación temporal de las innovaciones, en todos los casos se observa que el valor de  $\kappa$  aumenta con el nivel de agregación, contradiciendo el resultado esperado bajo el supuesto de autoafinidad de la distribución marginal de un proceso GARCH.

**Cuadro 3. Distribución MN(0,1; $\lambda,\pi$ ) (mezcla de normales)**

Variable Flujo		$(\lambda, \pi) = (0.50, 0.25)$		$(\lambda, \pi) = (0.50, 0.50)$		$(\lambda, \pi) = (0.50, 0.75)$	
		$\hat{\kappa}_T(m, \theta)$	$\hat{\kappa}_T(m, \hat{\theta}_T)$	$\hat{\kappa}_T(m, \theta)$	$\hat{\kappa}_T(m, \hat{\theta}_T)$	$\hat{\kappa}_T(m, \theta)$	$\hat{\kappa}_T(m, \hat{\theta}_T)$
T = 100	m = 1	12.665 (2.796)	7.264 (6.186)	12.717 (3.086)	7.910 (6.767)	12.524 (3.055)	6.839 (6.279)
	m = 2	14.258 (4.834)	8.259 (7.234)	14.178 (4.660)	7.818 (7.136)	13.973 (5.034)	7.529 (7.141)
T = 250	m = 1	11.632 (1.883)	9.462 (6.980)	11.709 (1.922)	9.270 (7.639)	11.598 (1.833)	8.977 (7.417)
	m = 2	12.171 (2.651)	9.664 (6.825)	12.264 (2.639)	8.786 (7.750)	12.503 (2.619)	8.523 (7.723)
	m = 5	14.039 (4.547)	9.248 (7.378)	14.096 (4.905)	8.519 (8.147)	13.857 (4.268)	8.386 (8.361)
T = 500	m = 1	11.213 (1.263)	10.957 (6.185)	11.202 (1.329)	9.499 (7.922)	11.205 (1.339)	9.794 (8.735)
	m = 2	11.578 (1.893)	10.607 (6.218)	11.631 (1.919)	9.418 (8.811)	11.651 (1.945)	9.353 (9.053)
	m = 5	12.568 (2.959)	11.068 (7.010)	12.800 (3.241)	9.330 (9.288)	12.809 (3.076)	8.438 (8.798)
	m = 10	14.149 (4.604)	11.273 (7.822)	13.853 (4.529)	7.968 (8.708)	14.038 (4.732)	8.318 (9.125)
T = 1000	m = 1	10.901 (0.931)	12.123 (5.156)	10.999 (0.992)	12.153 (8.967)	10.998 (0.945)	10.832 (9.194)
	m = 2	11.246 (1.347)	12.184 (5.834)	11.124 (1.229)	10.690 (9.221)	11.170 (1.286)	11.247 (9.446)
	m = 5	11.773 (2.013)	12.460 (6.289)	11.801 (2.055)	10.925 (9.535)	11.838 (2.065)	9.636 (9.929)
	m = 10	12.653 (3.096)	12.537 (7.056)	12.797 (3.123)	9.128 (9.768)	12.682 (2.946)	7.786 (9.435)
	m = 20	13.567 (4.271)	12.529 (7.497)	13.949 (4.491)	8.178 (9.899)	14.167 (4.652)	6.918 (9.617)

**Nota.** Mezcla de dos normales independientes:

$$z_t = (1 - \pi)z_{1,t} + \pi z_{2,t}, \begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} \sim N_2 \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \lambda\sigma^2 \end{pmatrix} \right\}, \sigma^2 = 1/((1 - \pi)^2 + \pi^2\lambda)$$

Para  $\lambda=0.5$ , al aumentar el valor de  $\pi$  aumenta la contribución del proceso  $z_1$  con mayor dispersión.

**Cuadro 4. Distribución Generalizada del Error, GED(v)**

Variable Flujo		GED(v=0.5)		GED(v=1.5)	
		$\hat{\kappa}_T(m, \theta)$	$\hat{\kappa}_T(m, \hat{\theta}_T)$	$\hat{\kappa}_T(m, \theta)$	$\hat{\kappa}_T(m, \hat{\theta}_T)$
T = 100	m = 1	4.083 (0.691)	3.403 (3.480)	13.508 (3.599)	7.991 (6.748)
	m = 2	5.310 (1.371)	3.628 (4.180)	14.622 (5.437)	8.182 (7.233)
T = 250	m = 1	3.629 (0.464)	2.609 (2.585)	12.215 (2.292)	9.776 (6.878)
	m = 2	4.434 (0.831)	2.949 (2.727)	12.786 (2.658)	10.286 (7.020)
	m = 5	6.651 (1.976)	3.511 (4.535)	14.109 (4.700)	10.164 (7.522)
T = 500	m = 1	3.473 (0.354)	2.144 (1.812)	11.515 (1.614)	11.412 (5.901)
	m = 2	4.067 (0.561)	2.071 (1.802)	11.965 (2.004)	12.077 (6.253)
	m = 5	5.487 (1.277)	2.157 (2.575)	12.910 (3.081)	11.234 (7.017)
	m = 10	7.638 (2.507)	2.365 (3.115)	13.799 (4.691)	11.628 (7.602)
T = 1000	m = 1	3.340 (0.273)	1.728 (1.441)	11.163 (1.139)	11.874 (4.583)
	m = 2	3.868 (0.457)	1.837 (1.001)	11.635 (1.567)	12.501 (5.124)
	m = 5	5.038 (0.947)	1.721 (1.713)	11.854 (2.104)	12.287 (5.060)
	m = 10	6.602 (1.643)	1.674 (2.385)	12.872 (3.113)	13.195 (6.358)
	m = 20	9.064 (3.284)	1.753 (2.805)	13.972 (4.594)	12.729 (6.701)

**Nota.** Si  $0 < v < 2$ , la distribución GED es de colas pesadas. Para  $v = 2$ , la distribución es normal, y de colas ligeras para  $v > 2$ .

**Cuadro 5. Innovaciones serialmente correlacionadas (GARCH semi-fuerte)**

Variable Flujo	Innovaciones serialmente correlacionadas (GARCH semi-fuerte)				
	Innovaciones AR(1) gaussianas: $z_t = \phi z_{t-1} + u_t$ ( $\phi = 0.15$ )		Innovaciones MA(1) gaussianas: $z_t = u_t - \phi u_{t-1}$ ( $\phi = 0.15$ )		
	$\hat{\kappa}_T(m, \theta)$	$\hat{\kappa}_T(m, \hat{\theta}_T)$	$\hat{\kappa}_T(m, \theta)$	$\hat{\kappa}_T(m, \hat{\theta}_T)$	
$T = 100$	$m = 1$	12.816 (3.174)	7.816 (6.511)	13.033 (3.198)	7.169 (6.014)
	$m = 2$	14.054 (4.798)	7.817 (6.972)	13.849 (4.414)	7.751 (6.918)
$T = 250$	$m = 1$	11.715 (1.868)	8.388 (5.699)	11.688 (1.856)	8.577 (5.833)
	$m = 2$	12.534 (2.606)	8.696 (5.747)	12.160 (2.629)	8.576 (5.948)
	$m = 5$	14.088 (4.782)	8.684 (6.405)	14.077 (4.742)	9.019 (6.289)
$T = 500$	$m = 1$	11.254 (1.295)	8.906 (4.582)	11.257 (1.286)	9.537 (4.3968)
	$m = 2$	11.664 (1.889)	9.738 (5.037)	11.747 (1.891)	9.556 (5.166)
	$m = 5$	12.719 (3.067)	10.553 (5.865)	12.715 (2.999)	10.536 (5.589)
	$m = 10$	13.808 (4.904)	10.330 (6.053)	14.019 (4.814)	10.814 (6.117)
$T = 1000$	$m = 1$	10.951 (0.942)	9.419 (3.565)	10.974 (0.945)	9.267 (3.565)
	$m = 2$	11.285 (1.319)	9.891 (3.834)	11.202 (1.284)	9.817 (3.449)
	$m = 5$	11.765 (2.012)	10.017 (3.959)	11.762 (2.059)	10.353 (4.109)
	$m = 10$	12.743 (3.129)	10.459 (4.824)	12.639 (3.047)	10.658 (4.454)
	$m = 20$	13.617 (4.638)	10.853 (5.017)	14.033 (4.893)	11.088 (5.394)

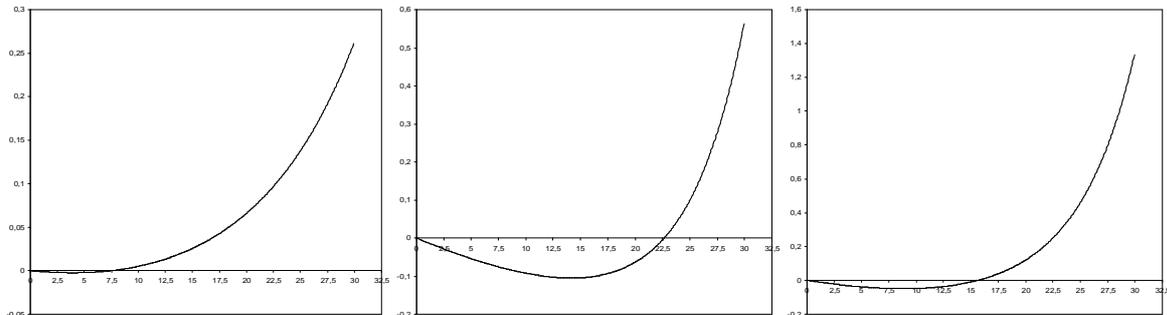
Bajo correlación débil de las innovaciones  $z_t$ , se observa un pequeño efecto de reducción del valor de  $\kappa$ , especialmente en el caso de  $\theta$  desconocido, que es prácticamente el mismo tanto en el caso AR(1) como MA(1).

## 6. Aplicación a series de rendimientos diarios de tipos de cambio

Finalmente, realizamos una aplicación del estimador IL (2007) a tres series de rendimientos diarios de tipos de cambio: \$/euro, yen/\$ y \$/libra esterlina para el período 03.01.2000-18.01-2008, con un total de  $T = 2025$  observaciones. En los tres casos no existe evidencia de dinámica regular en el nivel de la serie y la estimación QMV bajo normalidad condicional de un modelo GARCH(1,1) presenta los resultados habituales: elevada significación de los parámetros estimados y casi integración,  $\hat{\alpha}_{1,T} + \hat{\beta}_{1,T} \sim 1$ .

El siguiente gráfico 1 presenta el comportamiento de la función (4.2).

**Gráfico 1.** Función  $\Psi_T(\kappa, \hat{\theta}_T)$  para valores  $\kappa \in (0, \kappa_0]$  (de izquierda a derecha: TC1, TC2, TC3)



Los principales resultados se muestran en el siguiente cuadro 6. En el caso de agregación se presentan los resultados para el caso de variables flujo. Los resultados

para el caso de variables stock son casi idénticos.

**Cuadro 6. Resultados para las series de tipos de cambio**  
Caso de agregación con variables flujo

	TC1: \$/euro		TC2: yen/\$		TC3: \$/libra	
	$\hat{\kappa}_T(\hat{\theta}_T)^{(*)}$		$\hat{\kappa}_T(\hat{\theta}_T)^{(*)}$		$\hat{\kappa}_T(\hat{\theta}_T)^{(*)}$	
$m = 1$	7.51 (2.468)		22.60 (2.533)		15.66 (1.644)	
	$\hat{\kappa}_T(m, \hat{\theta}_T)$	$\hat{\kappa}_T(m, \hat{\theta}_T(m))$	$\hat{\kappa}_T(m, \hat{\theta}_T)$	$\hat{\kappa}_T(m, \hat{\theta}_T(m))$	$\hat{\kappa}_T(m, \hat{\theta}_T)$	$\hat{\kappa}_T(m, \hat{\theta}_T(m))$
$m = 2$	7.36	7.26	24.86	22.06	15.46	16.11
$m = 5$	8.66	8.71	34.11	32.81	17.71	22.86
$m = 10$	10.41	11.21	49.66	89.41	17.76	34.41
$m = 20$	12.41	15.46	76.66	-	17.26	-

(\*) Entre paréntesis, error estándar asintótico empleando el kernel de Bartlett con ancho de ventana fijo y proporcional al tamaño muestral para la estimación de la varianza a largo plazo (ver detalles en Iglesias y Linton (2007)).

En este cuadro,  $\hat{\kappa}_T(m, \hat{\theta}_T(m))$  denota el estimador de  $\kappa$  en el caso de agregación de las innovaciones estimadas empleando como valores de  $\theta$  los valores estimados corregidos por agregación temporal en el caso GARCH semi-fuerte como en Drost y Nijman (1993) a efectos de comparación con los resultados basados directamente en los valores estimados para la serie completa.

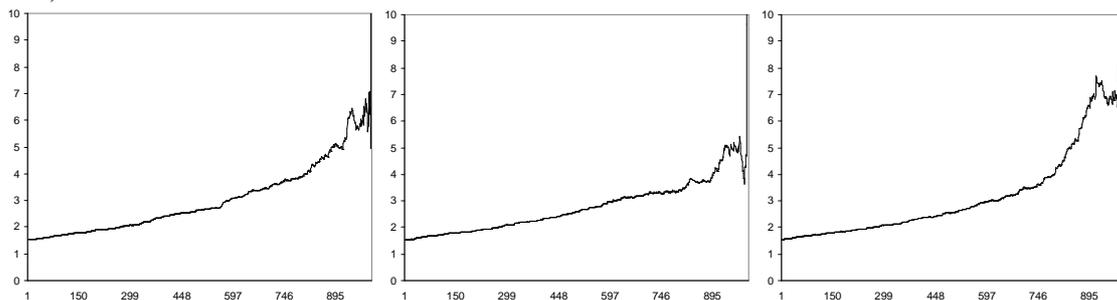
Salvo en el caso de la serie de tipos de cambio \$/euro, donde los resultados son muy similares bajo agregación, en el resto de casos los resultados difieren sustancialmente. Esto podría tomarse como evidencia de que para esta primera serie el proceso GARCH no se caracteriza por innovaciones *i.i.d.*.

A efectos de comparación presentamos los resultados del estimador de Hill (1975) en su versión más simple descrita en Quintos, Fan y Phillips (2001),

$$\hat{\alpha}_T(m_T) = \left( \frac{1}{m_T} \sum_{j=1}^{m_T} \ln(r_{T-j+1:T}^2) - \ln(r_{T-m_T:T}^2) \right)^{-1}$$

donde  $r_{T-j+1:T}^2$  denota el  $j$ -ésimo mayor valor de la serie de rendimientos del tipo de cambio. Teniendo en cuenta que  $\hat{\alpha}_T(m_T) \xrightarrow{p} \kappa/2$ , para hacer comparables los resultados, el gráfico 2 presenta  $\tilde{\alpha}_T(m_T) = 2\hat{\alpha}_T(m_T) \xrightarrow{p} \kappa$  para  $m_T = 1, \dots, 1000$ .

**Gráfico 2.** Estimador de Hill (1975) del índice de cola,  $\tilde{\alpha}_T(m_T) = 2\hat{\alpha}_T(m_T)$  (de izqda. a dcha: TC1, TC2, TC3)



## Referencias

- Basrak, B., R.A. Davis, T. Mikosch (2002). Regular Variation of GARCH Processes. *Stochastic Processes and Their Applications*, 99(1) (95-115)
- Berkes, I., L. Horváth (2003). The Rate of Consistency of the Quasi-Maximum Likelihood Estimator. *Statistics & Probability Letters*, 61 (133-143).
- Berkes, I., L. Horváth (2004). The Efficiency of the Estimators of the Parameters in GARCH processes. *The Annals of Statistics*, 32(2) (633-655).
- Berkes, I., L. Horváth, P. Kokoszka (2003). Estimation of the Maximal Moment Exponent of a GARCH(1,1) Sequence. *Econometric Theory*, 19 (565-586).
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, 31 (307-327).
- Carrasco, M., X. Chen (2002). Mixing and Moment Properties of various GARCH and Stochastic Volatility Models. *Econometric Theory*, 18 (17-39).
- Dahl, C.M., E.M. Iglesias (2007). Asymptotic Normality of the QMLE of Stationary and Nonstationary GARCH with serially dependent innovations. *Manuscript*. School of Economics and Management, University of Aarhus.
- de Haan, L., S.I. Resnick, H. Rootzén, C.G. de Vries (1989). Extremal Behaviour of Solutions to a Stochastic Difference Equation with Applications to ARCH Processes. *Stochastic Processes and their Applications*, 32 (213-224).
- Drost, F.C., T.E. Nijman (1993). Temporal Aggregation of GARCH Processes. *Econometrica*, 61(4) (909-927).
- Glosten, L.R., R. Jagannathan, D.E. Runkle (1993). On the Relationship Between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks. *Journal of Finance*, 48 (1779-1801).
- Hall, P., Q. Yao (2003). Inference in ARCH and GARCH models with Heavy-Tailed Errors. *Econometrica*, 71(1) (285-317).
- He, C., T. Teräsvirta (1999). Properties of Moments of a Family of GARCH Processes. *Journal of Econometrics*, 92 (173-192).
- Hill, B. (1975). A Simple General Approach to Inference about the Tail of a Distribution. *The Annals of Mathematical Statistics*, 3 (1163-1174).
- Huisman, R., K.G. Koedijk, C.J.M. Kool, F. Palm (2001). Tail-Index Estimates in Small Samples. *Journal of Business and Economic Statistics*, 19(1) (208-216).
- Iglesias, E.M., O.B. Linton (2007). Estimation of Tail Thickness Parameters from GARCH Models. *Workshop in Econometrics, Economics Department. Indiana University*.
- Kearns, P., A. Pagan (1997). Estimating the Density Tail Index for Financial Time Series. *Review of Economics and Statistics*, 79(2) (171-175).
- Kesten, H. (1973). Random Difference Equations and Renewal Theory for Products of Random Matrices. *Acta Mathematica*, 131 (207-248)
- Kulperber, R., H. Yu (2005). High Moment Partial Sum Processes of Residuals in GARCH Models and Their Applications. *The Annals of Statistics*, 33(5) (2395-2422).
- Lindner, A.M. (2007). Stationarity, Mixing, Distributional Properties and Moments of GARCH( $p,q$ ). En: Andersen, T.G., R.A. Davis, J.P. Kreib, T. Mikosch (eds.), *Handbook of Financial Time Series*, Springer.
- Ling, S., M. McAleer (2002). Stationarity and the Existence of Moments of a Family of GARCH Processes. *Journal of Econometrics*, 106 (109-117).
- Liu, J.C. (2006). On the Tail Behaviors of a Family of GARCH Processes. *Econometric Theory*, 22 (852-862).

- Mikosch, T., C. Stàricà (2000). Limit Theory for the Sample Autocorrelations and Extremes of a GARCH(1,1) Process. *The Annals of Statistics*, 28(5) (1427-1451).
- Mikosch, T., D. Straumann (2006). Stable Limits of Martingale Transforms with Application to the Estimation of GARCH Parameters. *The Annals of Statistics*, 34(1) (493-522).
- Nelson, D.B. (1990). Stationarity and Persistence in the GARCH(1,1) Model. *Econometric Theory*, 6 (318-334).
- Quintos, C., Z. Fan, P.C.B. Phillips (2001). Structural Change Tests in Tail Behaviour and the Asian Crisis. *Review of Economic Studies*, 68 (633-663).
- Stàricà, C., O. Pictet (1997). The Tales the Tails of GARCH(1,1) Processes Tell. *Unpublished manuscript*. University of Pennsylvania.
- Wagner, N., T.A. Marsh (2005). Measuring Tail Thickness under GARCH and an Application to Extreme Exchange Rate Changes. *Journal of Empirical Finance*, 12 (165-185).