INVESTIGACIONES ECONÓMICAS. vol. XXVI (3), 2002, 395-445

MODELOS DE MEMORIA LARGA PARA SERIES ECONÓMICAS Y FINANCIERAS

ANA PÉREZ
Universidad de Valladolid
ESTHER RUIZ
Universidad Carlos III de Madrid

En este trabajo se hace una revisión de los modelos de series temporales con memoria larga para la media y la varianza condicionada, con especial atención a los modelos ARMA fraccionalmente integrados (ARFIMA) y a los modelos GARCH y SV fraccionalmente integrados. Se estudian sus propiedades más importantes y se discute su aplicación en la modelización de series económicas y financieras. También se describen los principales métodos de estimación propuestos para estos modelos y se revisan algunos contrastes para detectar la presencia de memoria larga. Finalmente, se revisan los principales resultados sobre predicción de valores futuros de series temporales con memoria larga.

Palabras clave: Integración fraccional, persistencia, dominio de las frecuencias, densidad espectral, volatilidad.

(*JEL* C22)

1. Introducción

La propiedad de memoria larga suele relacionarse con la persistencia que muestran las autocorrelaciones muestrales de ciertas series temporales estacionarias, que decrecen a un ritmo muy lento, pero finalmente convergen hacia cero, indicando que las innovaciones de dichas series tienen efectos transitorios pero perduran durante mucho tiempo. Este comportamiento no es compatible ni con los modelos estacionarios que imponen un decrecimiento exponencial en las autocorrelaciones y, por lo tanto, en los efectos de las innovaciones, ni con el grado extremo de

Agradecemos a F. Marmol y C. Velasco la lectura atenta de parte de este trabajo. Sus comentarios y sugerencias nos han ayudado a clarificar muchas de las ideas aquí expuestas y han contribuido a mejorar la presentación final. Agradecemos asimismo la ayuda financiera de los proyectos SEC97-1379 (CICYT) y PB98-0026 (DGCICYT).

persistencia de los modelos integrados no estacionarios en los que las innovaciones tienen efectos permanentes.

La evidencia empírica sobre memoria larga se remonta a Hurst (1951) en el campo de la hidrología. Sin embargo, el interés por los modelos con memoria larga para series económicas surge a partir de los trabajos de Granger (1980) y Granger y Joyeux (1980), quienes observan que muchas de dichas series son aparentemente no estacionarias en media, y sin embargo, las series diferenciadas suelen presentar claros indicios de sobrediferenciación. Por lo tanto, para modelizar este tipo de series, la diferenciación parece "excesiva" pero la no diferenciación tampoco es adecuada. Para cubrir este vacío entre los casos extremos de modelos ARIMA con raíces unitarias, típicamente utilizados para modelizar series no estacionarias cuyo nivel evoluciona temporalmente, y modelos ARMA estacionarios donde el nivel medio es constante y la serie vuelve relativamente rápido a dicho nivel, Granger (1980), Granger y Joyeux (1980) y Hosking (1981) proponen una clase de procesos intermedios en los que el orden de integración es fraccional. Se trata de los procesos ARMA fraccionalmente integrados, ARFIMA(±[¥[²), donde ¥ es un número real. Al permitir que el orden de integración, ¥, sea un número no entero, estos modelos actúan como un "puente" entre los procesos con raíces unitarias ARIMA (¥ = 1) y los procesos ARMA estacionarios (¥ = 0). Como veremos más adelante, cuando 0 \ ¥ \ 1]2, los procesos ARFIMA son estacionarios, es decir, su nivel medio es constante, pero las desviaciones de la serie sobre dicho nivel tienen mayor duración que cuando ¥ = 0.

La presencia de memoria larga en series económicas, puede estar justificada por lo que Granger (1966) denominó la forma "típica" del espectro de las series económicas, que se caracteriza por no estar acotado en las frecuencias bajas y decrecer hiperbólicamente hacia cero. A esto hay que añadir los resultados sobre agregación de series de Robinson (1978) y Granger (1980, 1990), que demuestran que la suma de procesos AR(1) independientes cuyos coeficientes siguen una distribución tipo Beta (véase Johnson y Kotz, 1970, pp. 37-56), es un proceso fraccionalmente integrado. Dado que muchas variables económicas son agregados de otras variables, este resultado podría explicar la presencia de memoria larga en ciertas series económicas. Otra explicación alternativa de la existencia de memoria larga en los agregados económicos puede verse en Parke (1999).

La evidencia empírica sobre la presencia de memoria larga en series económicas y financieras es extensa. Por nombrar algunos ejemplos, Greene y Fielitz (1977) utilizan el estadístico de rango reescalado para contrastar la presencia de memoria larga en 200 series de rendimientos y encuentran evidencia en un número elevado de ellas. Posteriormente, Lo (1991) detecta memoria larga en rendimientos financieros utilizando una modificación del mencionado estadístico. También, Cheung (1993a) y Baillie y Bollerslev (1994) encuentran evidencia de memoria larga en activos financieros. En series macroeconómicas, por ejemplo, Diebold y Rudebush (1989) y Sowell (1992b) encuentran memoria larga en series trimestrales del Producto Nacional Bruto americano y Delgado y Robinson (1993), Hassler y Wolters (1995) y Baillie, Chung y Tieslau (1996) en distintas series mensuales de inflación. Por otra parte, la existencia de memoria larga estacional ha sido observada empíricamente, entre otros, por Porter-Hudak (1990), Ray (1993a) o Franses y Ooms (1997).

En los últimos años, ha surgido también un gran interés por la utilización de procesos con memoria larga para la modelización de la volatilidad de series financieras. El trabajo pionero de Ding, Granger y Engle (1993) puso de manifiesto que las autocorrelaciones muestrales de ciertas transformaciones de los rendimientos absolutos del índice bursátil S&P500 decrecen muy lentamente hacia cero, en consonancia con la propiedad de memoria larga. Trabajos posteriores, como Crato y de Lima (1994), Bollerslev y Mikkelsen (1996) y Lobato y Robinson (1998) han confirmado la evidencia de memoria larga en los cuadrados de distintas series financieras. Motivados por estos resultados, se han propuesto recientemente varios modelos que tratan de representar la propiedad de memoria larga en los momentos condicionados de segundo orden.

En la actualidad, existen varias revisiones de la literatura sobre modelos con memoria larga; véanse, por ejemplo, Beran (1994), Robinson (1994b) y Baillie (1996). El objetivo de este artículo es actualizar dichas revisiones incluyendo además, por primera vez, una revisión de los resultados relativos a predicción con modelos ARFIMA y a modelos con memoria larga en la varianza condicional. Aunque los modelos ARFIMA no son los únicos capaces de generar series con memoria larga, en este trabajo nos vamos a centrar en dichos modelos por ser los más utilizados en el análisis empírico. El artículo se organiza de la siguiente forma. En la sección 2 se describen las principales características de

los procesos ARFIMA tanto en el dominio del tiempo como en el dominio de las frecuencias. En la sección 3 se describen varios métodos de estimación y en la sección 4, contrastes de la hipótesis de memoria larga. En la sección 5 se revisa la literatura sobre predicción con modelos ARFIMA. La sección 6 revisa los principales modelos con memoria larga para la varianza condicionada. Finalmente, en la sección 7 se presentan las conclusiones. Los resultados descritos en este artículo son ilustrados con el análisis de varias series simuladas y reales.

2. Procesos con memoria larga: definición y propiedades

En esta sección se describen las principales propiedades estadísticas de series temporales con memoria larga, con especial atención a series generadas por modelos ARFIMA. La propiedad de memoria larga en una serie temporal suele entenderse como la existencia de una dependencia no despreciable entre observaciones que distan entre sí largos periodos de tiempo. Brockwell y Davies (1991) y Beran (1994) definen la memoria larga en un proceso estacionario $^{1}{}_{\mu}$ con función de autocorrelación $:(\neg)$y densidad espectral $\S(5)$, si existe un número real $0 \setminus + \setminus 1$ y una constante $\mathsf{c}_{:} \wedge 0$, tales que $:(\neg)\ 8\ \mathsf{c}_{:}|\neg|^{\mathrm{P}+}$ cuando $\neg\ \mathrm{A}\ \mathrm{Q}$, o equivalentemente, si existe un número real $0 \setminus , \setminus 1$ y una constante $\mathsf{c}_{\S} \wedge 0$ tales que $\S(5)\ 8\ \mathsf{c}_{\S}|5|^{\mathrm{P}}$, cuando $5\ \mathrm{A}\ 0^{+}$, donde el símbolo 8 significa que el cociente de los términos a su derecha e izquierda converge hacia uno. Por tanto, desde el punto de vista de las autocorrelaciones, la memoria larga implica que éstas tienden a cero muy lentamente y no son absolutamente sumables. De hecho, la propiedad de memoria larga sólo determina la tasa de convergencia hacia cero de las correlaciones, pero no la magnitud de cada una de ellas. Cada correlación individual puede ser arbitrariamente pequeña, pero el decrecimiento debe ser lento. En el caso de la función de densidad espectral, ésta contiene información sobre el peso relativo de cada frecuencia en la variabilidad total de la serie, siendo el periodo asociado a cada frecuencia 5 igual a $29\,]5$. Cuando $5 = 0$, el periodo correspondiente es infinito, y por tanto, en series con memoria larga, donde la densidad espectral no está acotada en la frecuencia $5 = 0$, el componente a largo plazo domina a cualquier otro componente de corto plazo. Es importante señalar que la definición de memoria larga es asintótica, en el sentido de que informa del comportamiento de las correlaciones cuando $\neg\ \mathrm{A}\ \mathrm{Q}$ o del espectro cuando $5\ \mathrm{A}\ 0^{+}$.

Aunque existen otras definiciones más generales de memoria larga, como, por ejemplo, las de Robinson (1994b) y Baillie (1996), la definición anterior es suficiente para los procesos ARFIMA(±[¥[²), que son los más utilizados en el análisis empírico de series económicas y financieras, y son los que vamos a considerar en este artículo. En general, se dice que un proceso ¹µ es integrado de orden ¥, y se denota por I(¥), si se verifica que (1 l)^¥ ¹µ = ¶µ, donde l es el operador retardo, l^a ¹µ = ¹µP^a, y ¶µ es un proceso estacionario con densidad espectral acotada en la frecuencia cero. Si ¥ = 0, ¹µ es estacionario; si ¥ = 1[¹µ tiene una raíz unitaria; si ¥ es otro número entero, ¹µ tiene ¥ raíces unitarias; y si ¥ es un número real no entero, ¹µ es un proceso fraccionalmente integrado. Si además ¶µ es un proceso ARMA(±[²) decimos que ¹µ es un proceso ARFIMA(±[¥[²). En el caso más sencillo, donde ± = ² = 0, no hay dependencia dinámica a corto plazo en la serie ¹µ, y el proceso resultante, conocido como ruido blanco fraccionalmente integrado o ARFIMA(0[¥[0), viene dado por la siguiente ecuación:

$$(1 \quad l)^{¥}\, ¹_µ = 1_µ \qquad [1]$$

donde $1_µ$ es un ruido blanco de media cero y varianza $;^2_1$. Las propiedades estadísticas del proceso ¹µ definido en [1] han sido analizadas por Hosking (1981). En particular, si 1]2 \ ¥ \ 1]2, ¹µ es estacionario e invertible, y admite la siguiente representación AR(Q):

$$(1 \quad l)^{¥}\, ¹_µ = \sum_{¬=0}^{?} 9_¬ \, ?_µP¬ = 1_µ \qquad [2]$$

donde $9_¬ = \frac{h(¬P¥)}{h(P¥)h(¬+1)}$, y por ser ¹µ estacionario, tiene además una representación MA(Q) dada por:

$$¹_µ = (1 \quad l)^{P¥} 1_µ = \sum_{¬=0}^{?})_¬ 1_µP¬ \qquad [3]$$

donde $)_¬ = \frac{h(¬+¥)}{h(¥)h(¬+1)}$, siendo (•) la función gamma (véase Brockwell y Davies, 1991, p. 521). Utilizando la aproximación [(¢ + ¹)] (£ + ¹)] 8 ¹^¢P£ cuando ¹ A Q, se comprueba que, cuando ¬ A Q,)_¬ 8 $\frac{¬^{¥P1}}{h(¥)}$, y por tanto, los coeficientes de la representación de Wold del proceso ARFIMA(0[¥[0) convergen hiperbólicamente hacia cero, en contraposición con el decrecimiento exponencial típico de los procesos ARMA. De hecho, dichos coeficientes satisfacen la condición de

estacionariedad $\sum_{\neg=0}^{?})\,\frac{2}{\neg} \setminus \mathrm{Q}$, pero no son absolutamente sumables, es decir, $\sum_{\neg=0}^{?} |\,)_{\neg}| = \mathrm{Q}$.

La propiedad más característica de los procesos ARFIMA estacionarios (¥ \ 1]2) es el comportamiento asintótico de su función de autocorrelación, o equivalentemente, el de su densidad espectral en las frecuencias bajas. Según la ecuación [3], $^{1}{}_{\mu}$ es el proceso resultante de un sistema lineal donde el proceso de entrada, 1_{μ}, es un ruido blanco con función de densidad espectral $\S_1(5) = ;\frac{2}{1}]29$. Por tanto, la densidad espectral de $^{1}{}_{\mu}$ se obtiene a partir de la de 1_{μ} de la siguiente forma (véase Priestley 1981, p. 266):

$$\S(5) = \S_1(5)\left|1 \quad {}^{\mathrm{a}}5\right|^{\mathrm{P}2¥} = \frac{;\frac{2}{1}}{29}\,[2(1 \quad ¤°´5)]^{\mathrm{P}¥} = \frac{;\frac{2}{1}}{29}\,[2´{}^{-}(5]2)]^{\mathrm{P}2¥}$$

y cuando 5 A 0^{+} se tiene la siguiente aproximación: $\S(5)$ 8 $\frac{;\frac{2}{1}}{29}5^{\mathrm{P}2¥}$. De esta expresión se deduce que para 0 \ ¥ \ 1]2, la densidad espectral de $^{1}{}_{\mu}$ es una función decreciente de 5 no acotada en el origen y el proceso $^{1}{}_{\mu}$ tiene memoria larga, según la definición de Robinson (1994b) y Baillie (1996). Por el contrario, cuando 1]2 \ ¥ \ 0, la densidad espectral se anula en el origen y está dominada por las frecuencias altas. En este caso, se dice que el proceso $^{1}{}_{\mu}$ presenta dependencia negativa o antipersistencia. Aunque el rango 1]2 \ ¥ \ 0 no tiene mucho interés en sí mismo, puede resultar útil en la modelización de las primeras diferencias de procesos ARFIMA(0[¥[0) no estacionarios con orden de integración 1]2 \ ¥ \ 1. Como veremos más adelante, estos procesos pueden modelizar el comportamiento de series no estacionarias pero que finalmente revierten hacia la media, propiedad que no pueden modelizar los procesos con raíces unitarias.

La función de autocovarianza del proceso ARFIMA(0[¥[0) viene dada por la siguiente expresión:

$$\text{-}(\neg) = ;\frac{2}{1}\frac{(1 \quad 2¥)\ (\neg + ¥)}{(¥)\ (1 \quad ¥)\ (\neg + 1 \quad ¥)}[\,\text{si } \neg \ 5\ 0[$$

a partir de la cual se obtiene la siguiente expresión de la función de autocorrelación:

$$:(\neg) = \frac{(1 \quad ¥)\ (\neg + ¥)}{(¥)\ (\neg + 1 \quad ¥)} = \prod_{1\mathrm{A}^{\mathrm{a}}\mathrm{A}\neg} \frac{^{\mathrm{a}} + ¥ \quad 1}{^{\mathrm{a}} \quad ¥}[\,\text{para } \neg \ 5\ 1$$

y aplicando nuevamente la aproximación [(¢ + ¹)] (£ + ¹)] 8 ¹¢P£, resulta que:

$$:(\neg)\ 8\ \frac{(1\quad ¥)}{(¥)}\neg^{2¥P1}[\ \text{cuando}\ \neg\ \mathrm{A}\ \mathrm{Q}$$

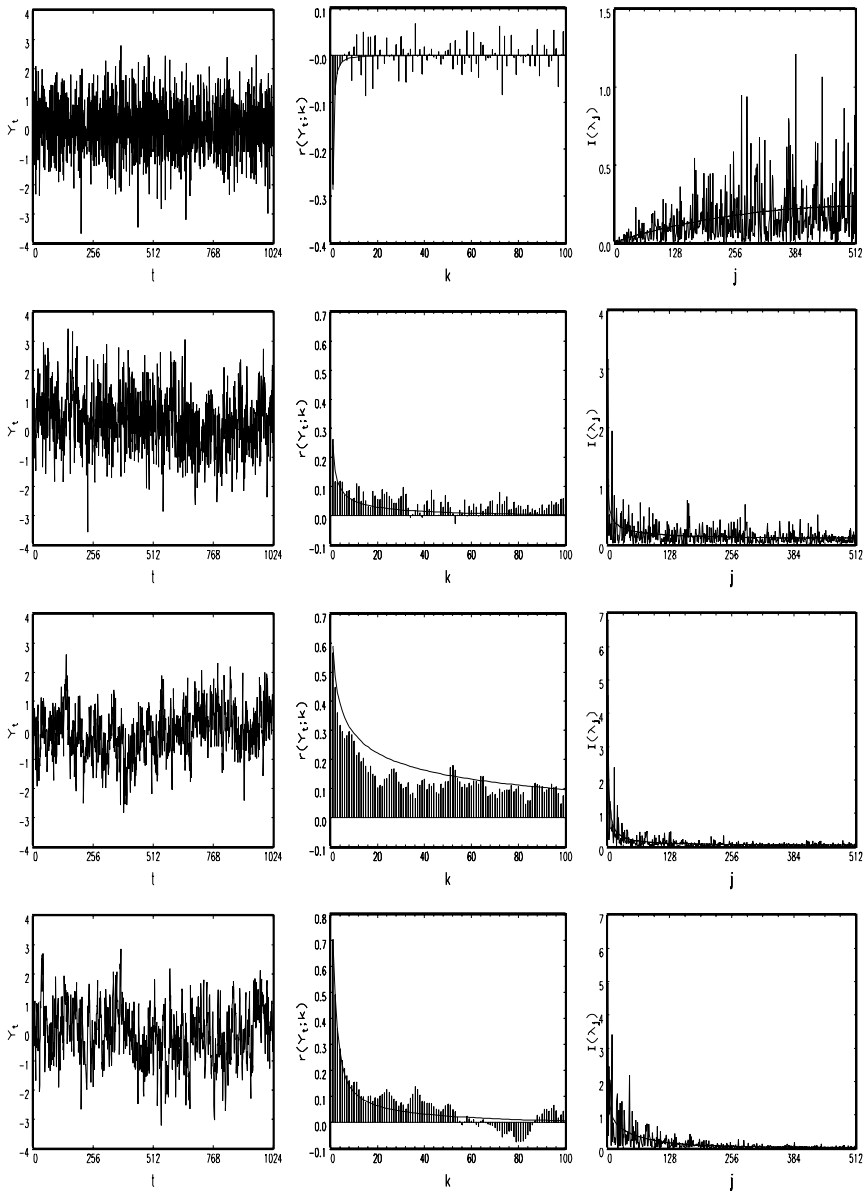
Esta última expresión indica que, cuando 0 \ ¥ \ 1]2, la función de autocorrelación del proceso ARFIMA(0[¥[0) es siempre positiva y decrece a un ritmo hiperbólico ¬²¥P¹, a diferencia de los procesos ARMA, cuya función de autocorrelación decrece exponencialmente. Para ilustrar esta dependencia a largo plazo típica de los procesos ARFIMA, en el Cuadro 1 se comparan las correlaciones de un proceso AR(1) de la forma $¹_µ = 0.5¹_{µP1} + 1_µ$ con las de un proceso ARFIMA(0,1/3,0). Aunque en ambos casos la primera autocorrelación es 0.5, el ritmo con que las correlaciones convergen hacia cero es muy diferente: mientras en el proceso AR(1) la autocorrelación de orden 25 es prácticamente cero, en el ARFIMA es 0.173 y sólo se reduce a 0.109 en el retardo 100.

CUADRO 1
Comparación de las funciones de autocorrelación de un proceso AR(1), $¹_µ = 0[5¹_{µP1} + 1_µ$ y un proceso ARFIMA, $(1\quad 1)^{1]3}¹_µ = 1_µ$

Retardo(*k*)	:(¬) [ARFIMA(0,1/3,0)]	:(¬) [AR(1) con > = 0[5]
1	0.500	0.500
2	0.400	0.250
3	0.350	0.125
4	0.318	0.063
5	0.295	0.031
10	0.235	0.001
25	0.173	$2.98\text{x}10^{-8}$
50	0.137	$8.88\text{x}10^{-16}$
100	0.109	$7.89\text{x}10^{-31}$

Para ilustrar las propiedades de los modelos ARFIMA descritas hasta el momento, en el Gráfico 1 se representan tres series generadas por procesos ARFIMA(0[¥[0) con ¥ = { 0.4, 0.2, 0.4}, respectivamente. Cada una de estas series se acompaña de su función de autocorrelación y de su función de densidad espectral, teórica y muestral. Las series se han generado utilizando el algoritmo propuesto por Davies y Harte (1987). La función de autocorrelación teórica aparece corregida por su sesgo, obtenido con las expresiones de Hosking (1996). Este gráfico muestra la persistencia en las correlaciones y la concentración del espectro en las frecuencias bajas, cuando ¥ ^ 0, y la antipersistencia, cuando ¥ \ 0. Además, se observa que la concentración del espectro

GRÁFICO 1
Series y_t generadas por procesos ARFIMA estacionarios con varianza uno junto con su función de autocorrelación, r(y_t;k), y su densidad espectral, I $(5_{\text{«}})$, muestral y poblacional. De arriba abajo, las tres primeras series representan procesos ARFIMA(0,*d*,0) con *d*=-0.4, *d*=0.2, *d*=0.4 y la última representa un ARFIMA(1,*d*,0) con $\{> = 0[5[¥ = 0[2\}$.



en torno a la frecuencia cero es mayor cuanto mayor y más próximo está ¥ a la frontera de la no estacionariedad (¥ = 1]2). En el Gráfico 1 puede observarse también la diferencia entre las funciones de autocorrelación de los procesos ARFIMA con ¥ \ 0 y ¥ ^ 0. Cuando ¥ =0.4, se aprecia claramente el decrecimiento lento de las correlaciones que, incluso en retardos alejados, son significativamente distintas de cero. Además, la persistencia de las correlaciones con ¥ =0.4 es mayor que con ¥ =0.2.

Cuando en el proceso ARFIMA(0[¥[0) que acabamos de describir, se incluyen componentes autorregresivas y de medias móviles, se tiene el proceso ARFIMA(±[¥[²), que admite una representación de la forma:

$$>(l)(1 \quad l)^{¥}\,{}^{1}{}_{µ} = 2(l)1_{µ} \qquad [4]$$

donde>(l)y 2(l) son polinomios de orden ± y ², respectivamente, cuyas raíces están fuera del círculo unidad. El proceso ARFIMA(±[¥[²) definido en [4] proporciona una gran flexibilidad en la modelización del comportamiento dinámico de series económicas, ya que permite describir simultáneamente las propiedades dinámicas en el largo plazo, a través del parámetro ¥, y la correlación en el corto plazo, a través de los parámetros de la parte ARMA del modelo. El comportamiento a largo plazo del proceso ARFIMA(±[¥[²) es similar al del proceso ARFIMA(0[¥[0), ya que para observaciones muy distantes, los efectos de los parámetros ARMA son prácticamente despreciables. En concreto, Hosking (1981) prueba que si ¥ \ 1]2 y todas las raíces de >(») = 0 están fuera del círculo unidad, ¹µ es estacionario, y si ¥ ^ 1]2 y todas las raíces de >(») = 0 están fuera del círculo unidad, ¹µ es invertible. En el caso estacionario (¥ \ 1]2), la función de densidad espectral del proceso ARFIMA(±[¥[²) es de la forma:

$$§(5) = \frac{;^2_1}{29}\left|\frac{2(¦^{a}5)}{>(¦^{a}5)}\right|^2 [2´¦^{-}(5]2)]^{P2¥} \qquad [5]$$

En consecuencia, si 0 \ ¥ \ 1]2, el espectro de ¹µ no está acotado en las frecuencias bajas y el proceso ¹µ tiene memoria larga. Además, en este caso, las autocorrelaciones decrecen hiperbólicamente y no son sumables. Las expresiones de dichas autocorrelaciones se derivan en Hosking (1981). Sin embargo, su cálculo se simplifica considerablemente utilizando la formulación alternativa de Sowell (1992a) y Chung (1994).

En el último panel del Gráfico 1 se representa una serie simulada de un proceso ARFIMA(1[¥[0) con {> =0.5[¥ =0.2} junto con su función

de autocorrelación y su densidad espectral, muestral y poblacional. En este gráfico se observa la concentración del espectro en torno al origen propia de los modelos con memoria larga, así como la presencia de correlaciones significativas incluso en retardos alejados. Por otra parte, se aprecia la influencia en el corto plazo del parámetro autorregresivo, que hace que la función de autocorrelación presente valores grandes en los primeros retardos y decrezca luego a un ritmo más rápido que el de los procesos ARFIMA(0[¥[0).

Alternativamente, las diferencias entre los modelos ARFIMA, los modelos ARMA de memoria corta y los modelos integrados ARIMA, pueden establecerse analizando la forma en que la serie $^{1}{}_{\mu}$ responde a un cambio unitario en la perturbación del modelo. Mientras en un modelo ARIMA, el efecto sobre el nivel futuro de la serie de un *shock* sobre la perturbación, es permanente y no hay reversión a la media, en un ARFIMA dicho efecto finalmente desaparece, aunque a un ritmo más lento que el exponencial de los procesos ARMA. Para medir el grado de persistencia de los *shocks* sobre la media, Diebold y Rudebusch (1989) y Haubrich y Lo (1989) proponen utilizar la función respuesta al impulso acumulativa. Los valores de dicha función para una serie $^{1}{}_{\mu}$ son los coeficientes del polinomio $¤(l) = 1 + ¤_1 l + ¤_2 l^2 + ZZZ$ tal que $^{1}{}_{\mu} = 0 + ¤(l)1_{\mu}$. El coeficiente $¤_¬$ mide el impacto que tiene una innovación unitaria en el momento µ sobre el nivel de la serie en µ+¬, y su límite, $¤_? = \lim_{¬ Y ?} {}^{a}® ¤_¬$, mide el impacto a largo plazo de dicha innovación. Si la serie $^{1}{}_{\mu}$ en primeras diferencias admite una representación de la forma $(1 \quad l)^{1}{}_{\mu} = ¢(l)1_{\mu}$ donde $¢(l) = 1 + ¢_1 l + ¢_2 l^2 + ZZZ[$ se prueba fácilmente que $¤_¬ = 1 + ¢_1 + ZZZ + ¢_¬$. Por tanto, el impacto a largo plazo que tiene una innovación unitaria sobre el nivel de la serie será $¤_? = 1 + \sum_{a=1}^{?} +_a = ¢(1)$. En el modelo ARFIMA(±[¥[²) definido en la ecuación [4], el polinomio $¢(l)$ es de la forma $¢(l) = (1 \quad l)^{1P¥} >(l)^{P1} 2(l)$, y utilizando los resultados de Gradshteyn y Ryzhik (1980, sec. 9.1) resulta que:

$$+(1) = f(¥ \quad 1[1[1;1) >(1)^{P1} 2(1) = \begin{cases} 0 & ´^{a} \quad 0 \setminus ¥ \setminus 1 \\ >(1)^{P1} 2(1) & ´^{a} \quad ¥ = 1 \end{cases}$$

donde $f(•[•[•;•)$ es la función hipergeométrica y $>(1)$ y $2(1)$ son las sumas de los coeficientes de los polinomios $>(l)$ y $2(l)$, respectivamente. Esto significa que en los procesos I(1), el efecto de un *shock* persiste indefinidamente, ya que $¢(1) \neq 0$, y por tanto, no hay rever-

sión a la media, mientras que en un proceso I(¥) con 0 \ ¥ \ 1, el efecto de un shock acaba desapareciendo (¢(1) = 0) y la serie ¹µ revierte finalmente a su media, aunque lo hará a un ritmo muy lento caracterizado, precisamente, por el parámetro ¥. Obsérvese que esto es cierto no sólo en el caso estacionario (0 \ ¥ \ 1]2), sino también en el intervalo 1]2 4 ¥ \ 1, donde el proceso no es estacionario. En este sentido, suele decirse que el ARFIMA no estacionario es "menos no estacionario" que los procesos I(1).

Los modelos ARFIMA han sido también extendidos de forma que puedan representar memoria larga estacional. Por ejemplo, Gray, Zhang y Woodward (1989) proponen una clase de modelos fraccionales generalizados mediante polinomios de Gegenbauer, que permiten modelizar series con un comportamiento periódico (cíclico o estacional) persistente en el largo plazo. Porter-Hudak (1990) propone un modelo estacional fraccionalmente integrado y aplica dicho modelo a distintas series de agregados monetarios. Ray (1993a) utiliza un modelo con memoria larga estacional para representar el comportamiento dinámico de rendimientos de un activo financiero. Finalmente, Franses y Ooms (1997) proponen un modelo en el que el parámetro de integración fraccional puede cambiar con la estación y lo utilizan para modelizar una serie de inflación trimestral del Reino Unido.

3. Estimación de procesos ARFIMA

Los métodos de estimación de modelos con memoria larga propuestos hasta el momento en la literatura pueden clasificarse en dos grandes grupos: métodos paramétricos, en los que se conoce de forma explícita toda la estructura de correlación o del espectro, tanto en el corto como en el largo plazo, y métodos semiparamétricos, en los que no se especifica un modelo completo de la estructura de correlación (o espectral) sino sólo su comportamiento en los retardos alejados (o en las frecuencias bajas)[1]. El enfoque paramétrico tiene la ventaja de que permite caracterizar todas las autocorrelaciones y no sólo su decrecimiento hiperbólico. Además, los métodos propuestos son métodos de Máxima Verosimilitud (MV) y por tanto, en condiciones bastante generales, tienen buenas propiedades asintóticas. Sin embargo, si el modelo pa-

[1] Otros autores distinguen entre métodos de estimación en dos etapas, en los que se estima en primer lugar el parámetro ¥ y en una segunda etapa los parámetros ARMA del modelo, y métodos de estimación en una etapa, que estiman conjuntamente, por máxima verosimilitud, todos los parámetros del modelo.

ramétrico que se supone está mal especificado, los estimadores serán inconsistentes y los métodos semiparamétricos pueden ser más apropiados, ya que son más robustos a una mala especificación del modelo. Los estimadores semiparamétricos también tienen buenas propiedades asintóticas, aunque su tasa de convergencia es inferior a la típica $\sqrt{\mathrm{t}}$ de los estimadores paramétricos.

Alternativamente, existen otros métodos de estimación que son menos utilizados en aplicaciones empíricas. Li y McLeod (1986), por ejemplo, estudian los estimadores máximo verosímiles del modelo ARFIMA aproximado que se obtendría truncando en un retardo suficientemente grande su representación AR(Q) en [2]. Chung y Baillie (1993) analizan las propiedades en muestras finitas de estimadores de mínimos cuadrados condicionales en el dominio temporal. Una breve descripción de ambos métodos puede encontrarse en Beran (1994) y Baillie (1996), respectivamente.

3.1. Métodos de estimación paramétricos

Con respecto a los métodos de estimación paramétricos, Yajima (1985) analiza las propiedades del estimador MV de los parámetros del proceso ARFIMA(0[¥[0) gaussiano, que se obtiene maximizando la función de verosimilitud

$$\mathrm{l}(2;x) = \quad \frac{\mathrm{t}}{2}\text{-}^{-}(29) \quad \frac{1}{2}\text{-}^{-}|\&(2)| \quad \frac{1}{2}\imath'\&(2)^{\mathrm{P}1}\imath \qquad [6]$$

donde $2 = (;{}^{2}_{1}[¥)$, t es el tamaño muestral, $x = ({}^{1}{}_{1}[\mathit{lll}[{}^{1}{}_{\mathrm{t}})'$ es el vector de observaciones y $\&(2)$ es la matriz de varianzas y covarianzas de ^{1}, demostrando que dicho estimador es consistente y asintóticamente normal. Dahlhaus (1989) extiende estos resultados a procesos ARFIMA(±[¥[²) gaussianos con media 6 desconocida. En este caso, el parámetro 6 en la función de verosimilitud, $\mathrm{l}(2[6;{}^{1})$, puede sustituirse por un estimador consistente, como la media muestral, ${}^{\mathrm{r}} = \frac{1}{\mathrm{t}}\sum_{\mu=1}^{\mathrm{t}} {}^{1}{}_{\mu}$, y el estimador MV de 2 se obtiene maximizando la función de verosimilitud resultante, $\mathrm{l}(2[{}^{\mathrm{r}};{}^{1})$. Dahlhaus (1989) demuestra, además, la eficiencia asintótica del estimador MV. Posteriormente, Sowell (1992a) proporciona un algoritmo recursivo para calcular la matriz de varianzas y covarianzas $\&(2)$ y obtener la función de verosimilitud gaussiana en [6].

Sin embargo, el estimador MV exacto es computacionalmente muy costoso. Para maximizar la función de verosimilitud hay que emplear

métodos numéricos que requieren calcular, en cada iteración, la inversa de la matriz &(2). Además, si ¥ está próximo a 1/2, la evaluación de la inversa de &(2) puede ser muy inestable, ya que las covarianzas apenas cambian y dicha matriz llega a ser casi singular. Recientemente, Chan y Palma (1998) han propuesto una forma más eficiente de calcular la función de verosimilitud en modelos ARFIMA gaussianos, utilizando la descomposición en los errores de predicción, obtenida de forma recursiva en un número finito de pasos a través del filtro de Kalman, evitando así invertir la matriz &(2). Además, estos autores proponen otro método de estimación que consiste en truncar la representación MA(Q) del modelo ARFIMA en primeras diferencias y maximizar la función de verosimilitud de este modelo truncado, utilizando su representación en el espacio de los estados. Este método simplifica considerablemente los cálculos y proporciona estimadores consistentes y asintóticamente normales y eficientes.

Como alternativa a los métodos de máxima verosimilitud exacta, se han propuesto otros métodos basados en maximizar una aproximación a la función de verosimilitud gaussiana en el dominio de las frecuencias. Fox y Taqqu (1986) consideran los estimadores MV aproximados de Whittle (1962), que se obtienen minimizando la siguiente función:

$$\mathrm{l}_{\mathrm{w}}(2;{}^{1}) = \frac{1}{49}\int_{\mathrm{P}9}^{9}\left[\text{-}^{-}\,\S(5;2) + \frac{\mathrm{i}_{\mathrm{t}}(5)}{\S(5;2)}\right]¥5 \qquad [7]$$

donde $\S(5;2)$ es la función de densidad espectral de ${}^{1}{}_{\mu}$ e $\mathrm{i}_{\mathrm{t}}(5)$ es el periodograma de $({}^{1}{}_{1}[ZZZ[\,{}^{1}{}_{\mathrm{t}})$, definido como:

$$\mathrm{i}_{\mathrm{t}}(5) = \frac{1}{29\,\mathrm{t}}\left|\sum_{\mu=1}^{\mathrm{t}}({}^{1}{}_{\mu}\quad {}^{\mathrm{r}})\,{}^{\mathrm{a}\mu 5}\right|^{2}$$

Fox y Taqqu obtienen resultados asintóticos para ciertos parámetros del modelo y bajo ciertas restricciones sobre la densidad espectral de ${}^{1}{}_{\mu}$. Dahlhaus (1989) extiende dichos resultados y demuestra que el estimador de Whittle del vector paramétrico completo, $\widehat{2}$, es consistente y asintóticamente normal en procesos ARFIMA estacionarios gaussianos y establece además condiciones para que dicho estimador sea asintóticamente eficiente.

Aunque los estimadores aproximados de Whittle son atractivos porque reemplazan la evaluación de la inversa de &(2) por transformadas de Fourier del vector de observaciones, tampoco son sencillos, desde un

punto de vista computacional, porque la integral en [7] debe evaluarse en cada iteración por métodos numéricos. La solución adoptada generalmente en estos casos es sustituir dicha integral por una suma finita sobre las frecuencias de Fourier y minimizar la función correspondiente:

$$\tilde{I}(2;\,^1) = \frac{1}{2t}\sum_{«=1}^{t\ P1}\left[\text{-}^{-}\,\S(5_«;2) + \frac{i_t(5_«)}{\S(5_«;2)}\right] \qquad [8]$$

donde $5_«$ es la «-ésima frecuencia de Fourier, $5_« = \frac{29«}{t}$. Minimizando $\tilde{I}(2;\,^1)$ con respecto a 2 se obtiene la versión discreta del estimador MV de Whittle en el dominio de las frecuencias, $\tilde{2}$. Dahlhaus (1989) demuestra que $\tilde{2}$ tiene las mismas propiedades asintóticas que $\widehat{2}$. Además, desde un punto de vista computacional, $\tilde{2}$ es un estimador más sencillo, porque el cálculo de $\tilde{I}(2;\,^1)$ se puede agilizar utilizando la simetría de la función de densidad espectral y del periodograma, y aplicando la transformada rápida de Fourier; Bloomfield (1976, cap. 4) o Priestley (1981, pp. 575-586) describen la transformada rápida de Fourier y sus aplicaciones.

Cheung y Diebold (1994) comparan las propiedades en muestras finitas de $\tilde{2}$ y del estimador MV exacto para procesos ARFIMA(0[¥[0) gaussianos y llegan a las siguientes conclusiones. Si la media del proceso es conocida, el estimador MV exacto es, por supuesto, más eficiente, pero cuando la media debe estimarse, el estimador $\tilde{2}$ es preferible, especialmente en muestras grandes donde calcular el estimador MV exacto resulta impracticable. Sin embargo, Smith, Sowell y Zin (1997) proponen una corrección por sesgo del estimador MV exacto que parece tener buenas propiedades en muestras finitas frente a $\tilde{2}$.

En el caso de procesos no gaussianos con memoria larga, Giriatis y Surgalis (1990) derivan el primer teorema central del límite para el estimador de Whittle bajo condiciones restrictivas sobre la densidad espectral del proceso. Hosoya (1996, 1997) presenta resultados más generales y deriva las propiedades asintóticas de los estimadores de Whittle para procesos no necesariamente gaussianos. Finalmente, Velasco y Robinson (2000) analizan las propiedades asintóticas y en muestras finitas del estimador de Whittle para procesos con memoria larga no estacionarios.

Como ilustración de los métodos de estimación MV que acabamos de describir, el Cuadro 2 recoge los resultados de la estimación de un modelo ARFIMA(0[¥[0) para la serie de inflación mensual en España,

observada desde enero de 1961 hasta diciembre de 2000, con un total de 480 observaciones. Las tasas de inflación se han obtenido como $^{1}{}_{\mu} = 100[\text{-}^{-}(\mathrm{p}_{\mu}) \quad \text{-}^{-}(\mathrm{p}_{\mu \mathrm{P}\, 1})]$, donde p_{μ} es el valor del Indice de Precios al Consumo (IPC) en el mes t. Los estimadores utilizados son el MV exacto, calculado siguiendo la propuesta de Chan y Palma (1998), y la versión discreta del estimador MV de Whittle en el dominio de las frecuencias. Los valores que aparecen entre paréntesis son las desviaciones típicas de las correspondientes distribuciones asintóticas de dichos estimadores. Ambos métodos de estimación evidencian la existencia de memoria larga en la serie mensual de inflación, aunque el estimador MV exacto proporciona un valor estimado del parámetro de integración fraccional superior al obtenido por el método de Whittle. Téngase en cuenta que, en muestras finitas, ambos estimadores presentan un sesgo negativo, que suele ser mayor en el estimador de Whittle que en el estimador MV exacto, especialmente cuando ¥ y t son pequeños; véase Cheung y Diebold (1994).

CUADRO 2
Estimación del parámetro d obtenida con distintos métodos para la serie mensual de inflación en España desde enero 1961 hasta diciembre 2000.

MV exacto	MV Whittle	GPH (1983) $m=T^{0.5}$	GPH (1983) $m=T^{0.6}$	GPH (1983) $m=T^{0.7}$	Robinson (1995a) con $\ell=1$ $m=T^{0.5}$	Robinson (1995a) con $\ell=1$ $m=T^{0.6}$	Robinson (1995a) con $\ell=1$ $m=T^{0.7}$
0.325	0.269	0.617	0.457	0.312	0.427	0.325	0.221
(0.036)	(0.036)	(0.137)	(0.101)	(0.074)	(0.137)	(0.101)	(0.074)

3.2. Métodos de estimación semiparamétricos

Alternativamente, los parámetros de modelos ARFIMA(±[¥[²) pueden estimarse por métodos semiparamétricos. Geweke y Porter-Hudak (1983) proponen un estimador, conocido como GPH, basado en un modelo de regresión para el logaritmo del periodograma. Este método explota la sencillez que tiene la expresión de la densidad espectral de los procesos ARFIMA(±[¥[²) en [5] en las frecuencias bajas. Evaluando dicha expresión en las frecuencias de Fourier y tomando logaritmos, se obtiene la siguiente aproximación:

$$\text{-}^{-}[\mathrm{i}_{\mathrm{t}}(5_{\ll})] \underline{\mathrm{8}} \ ,_{0} \quad ¥\left\{\text{-}^{-}\left[4\acute{}\,\text{¦}^{-2}(5_{\ll}]2)\right]\right\} + \mathrm{B}_{\ll}[\text{ con } \ll = 1[\,2[\,\mathbb{ZZ}[\ \mathrm{®} \quad [9]$$

donde $\mathrm{i}_{\mathrm{t}}(5_{\ll})$ es la «-ésima ordenada del periodograma, $,_{0} = \text{-}^{-}[\S_{\mathrm{u}}(0)]$ -, siendo - la constante de Euler (véase Abramovitz y Stegun, 1970, p. 255) y $\S_{\mathrm{u}}(5)$ la densidad espectral del proceso ARMA correspondiente, $\mathrm{B}_{\ll} = \text{-}^{-}[\mathrm{i}_{\mathrm{t}}(5_{\ll})]\S(5_{\ll})] + \text{-}$ y $\mathrm{®} = ¨(\mathrm{t})$ es un entero menor

que t y tendiendo hacia infinito más despacio que t . La expresión [9] puede interpretarse como un modelo de regresión en el dominio espectral, lo que sugiere estimar ¥ por mínimos cuadrados ordinarios (MCO). Geweke y Porter-Hudak (1983) argumentan que las perturbaciones B« de la ecuación [9] son asintóticamente independientes e idénticamente distribuidas con media cero y varianza 9^2]6, y establecen de forma heurística la consistencia y la normalidad asintótica del estimador, cuando ¥ \ 0. Porter-Hudak (1990) propone una extensión de este estimador para modelos con integración fraccional estacional. Posteriormente, Hassler (1993), Chen, Abraham y Peiris (1994) y Reisen (1994) proponen una modificación del estimador GPH utilizando en la regresión [9] el periodograma suavizado, en lugar del periodograma. Estos autores comprueban, mediante simulaciones, que estas modificaciones proporcionan estimaciones de ¥ menos sesgadas y con menor varianza.

Las propiedades asintóticas del estimador GPH no se establecen hasta el trabajo de Robinson (1995a), que propone una versión modificada más eficiente del mismo, y demuestra la consistencia y la normalidad asintótica para todo el rango de valores del parámetro 1]2 \ ¥ \ 1]2. La propuesta de Robinson es más general que GPH en varios sentidos. En primer lugar, sólo requiere que la densidad espectral en las frecuencias bajas verifique la condición:

$$\S(5)\ 8\ c_{\S}\ |5|^{P\,2¥}[\ \text{cuando}\ 5\ A\ 0^{+} \qquad [10]$$

sin explicitar la forma concreta de §(5); por tanto, el modelo de regresión a estimar es:

$$-^{-}[i_t(5_«)] \underline{8} -^{-}(c_{\S})\ \ 2¥(-^{-}|5_«|)+-^{-}[i_t(5_«)]\S(5_«)[\ \text{con}\ « = 1[\ 2[\ ZZZ[\ ®Z \qquad [11]$$

Robinson (1995a) sugiere además eliminar las frecuencias más bajas, truncando inferiormente en un punto ¡ que depende de t , de forma que el modelo [11] se estima en las frecuencias « = ¡ + 1[¡ + 2[ZZZ[®. Bajo hipótesis de normalidad para $^{1}{}_{µ}$, el estimador MCO de la ecuación [11] es consistente y asintóticamente normal, aunque con una tasa de convergencia $\overline{\pm}\overline{®}$ inferior a la $\overline{\pm}\overline{t}$ típica de los estimadores paramétricos gaussianos. Hurvich y Beltrao (1994) comprueban que, en muestras finitas, la reducción de sesgo que se consigue truncando inferiormente en ¡ no compensa el incremento en la varianza. Finalmente, Hurvich, Deo y Brodsky (1998) demuestran la consistencia del estimador GPH ori-

ginal y derivan expresiones del sesgo y la varianza de dicho estimador y establecen su distribución asintótica bajo hipótesis de normalidad.

El método de estimación GPH es muy sencillo de aplicar y no requiere apenas supuestos sobre el modelo. De hecho, el estimador GPH ha sido durante años el más utilizado en aplicaciones empíricas, al menos como punto de comparación para analizar las propiedades de otros estimadores. Sin embargo, la pérdida de eficiencia de este estimador puede llegar a ser extrema en muestras pequeñas. Agiakloglou, Newbold y Wohar (1993), por ejemplo, observan que cuando existe un componente AR(1) o MA(1) con parámetro cercano a la unidad, el estimador GPH tiene un sesgo enorme y es muy ineficiente. Otro problema adicional es la elección del número de frecuencias que se utilizan en la regresión. Pero quizá el mayor inconveniente de este método, según Robinson (1995b), es el carácter restrictivo del supuesto de normalidad para $^{1}{}_{\mu}$. En dos trabajos recientes, Velasco (1999b, 2000) extiende los resultados de Robinson (1995a) a modelos no estacionarios (¥ 5 1]2) y no necesariamente gaussianos, respectivamente.

Smith, Taylor y Yadav (1997) comparan las propiedades en muestras finitas del estimador GPH modificado de Robinson (1995a) con el estimador máximo verosímil exacto y el estimador semiparamétrico APER (Averaged Periodogram) de Lobato y Robinson (1996), y como era de esperar, concluyen que, en el supuesto de una correcta especificación del modelo, el estimador máximo verosímil exacto supera a los otros dos, proporcionando siempre estimaciones menos sesgadas y con menor varianza.

Alternativamente, Robinson (1995b) propone otro estimador semiparamétrico para procesos que verifiquen la condición [10]. Este estimador se obtiene maximizando una versión de la aproximación discreta de Whittle a la función de verosimilitud gaussiana en [8], donde la suma sólo se realiza en un entorno de la frecuencia cero. Este estimador domina a los anteriores en varios aspectos: es más eficiente, no impone normalidad sobre $^{1}{}_{\mu}$ y evita el recorte de las componentes en las frecuencias bajas así como la elección del número ® de ordenadas a promediar. Bajo ciertas condiciones, Robinson (1995b) prueba que el estimador es consistente y asintóticamente normal en el intervalo 1]2 \ $\pm\frac{¥}{\sqrt{t}}$ \ 1]2, aunque nuevamente la tasa de convergencia es inferior a $\sqrt{t}$. Velasco (1999a) extiende estos resultados al caso no estacionario (¥ 5 1]2).

Finalmente, Moulines y Soulier (1999, 2000) y Hurvich y Brodsky (2001) han propuesto, de forma independiente, otro estimador semiparamétrico para procesos gaussianos con memoria larga, que se obtiene estimando por MCO un modelo fraccional exponencial (FEXP) para el logaritmo del periodograma sobre todas las frecuencias de Fourier en el intervalo [0[9]. Bajo ciertas condiciones sobre la densidad espectral, estos autores demuestran que el estimador que proponen es consistente y obtienen expresiones de su sesgo y su varianza asintótica. Además, comprueban que, en ciertos casos, este estimador proporciona menor error cuadrático medio que el estimador GPH y que el estimador gaussiano semiparamétrico de Robinson (1995b).

Para ilustrar la aplicación de los métodos de estimación semiparamétricos, hemos incluido en el Cuadro 2 los resultados de la estimación GPH del parámetro ¥ para la serie de tasas mensuales de inflación analizada en la sección anterior. Como es habitual, hemos considerado diferentes valores del parámetro ® = ¨(t) que especifica el número de ordenadas a promediar, y para cada uno de estos valores, hemos omitido además las ¡ frecuencias más bajas, como sugiere Robinson (1995a). Las estimaciones del parámetro ¥ se acompañan de los valores del error estándar obtenido de la distribución asintótica de dichos estimadores. A la vista de los resultados, podemos concluir, en primer lugar, que la estimación de ¥ parece bastante inestable para los distintos valores de ®. Obsérvese también que al aumentar ®, se reduce la varianza de los estimadores y disminuye el valor estimado del parámetro ¥, como apuntan Hurvich y Beltrao (1994). Además, para los valores de ® más pequeños, los estimadores semiparamétricos proporcionan estimaciones de ¥ mayores que los estimadores máximo verosímiles, en consonancia con los resultados de Hassler y Wolters (1995). Finalmente, nótese que la varianza de los estimadores semiparamétricos es siempre mucho mayor que la de los estimadores MV.

4. Contrastes de hipótesis

En esta sección se presentan los métodos que mayor atención han recibido en la literatura econométrica para contrastar la presencia de memoria larga en una serie. También se describen brevemente algunos trabajos que analizan el comportamiento de los contrastes de raíces unitarias en el ámbito de los modelos con memoria larga. Finalmente, se consideran las propiedades de las autocorrelaciones muestrales en

presencia de memoria larga, por constituir éstas una de las herramientas más utilizadas en la detección de la memoria larga.

El contraste de Geweke y Porter-Hudak (1983) se deriva de la estimación MCO del modelo de regresión en el dominio espectral, definido en la ecuación [9]. La pendiente de este modelo es precisamente ¥, y por tanto, para contrastar la hipótesis $h_0 : ¥ = 0$ (memoria corta) frente a la alternativa $h_1 : ¥ \neq 0$, se realiza el contraste habitual de significación sobre el parámetro ¥ con una distribución µ-Student. Aunque este contraste es muy sencillo de aplicar, si existen bajo h_0 componentes autorregresivos o de medias móviles con valores grandes de los parámetros, la hipótesis nula se rechazará, aún siendo cierta, con mucha frecuencia, debido a que en estos casos, el estimador GPH de ¥ es muy sesgado; ver Agiakloglou, Newbold y Wohar (1993). Esta misma conclusión se deriva de las simulaciones de Cheung (1993b), que muestran un sesgo del contraste hacia alternativas del tipo ¥ ∧ 0 cuando el proceso generador de los datos es un AR(1) con parámetro autorregresivo grande.

Haubrich y Lo (1989) y Lo (1991) proponen un contraste de memoria corta basado en una modificación del estadístico r] s (rango reescalado) introducido por Hurst (1951). El estadístico r] s es el rango de las sumas parciales de las desviaciones de una serie a su media, normalizado por su desviación típica muestral. Lo (1991) propone reemplazar la desviación típica muestral por otro factor de normalización que recoja la posible correlación en el corto plazo que pudiera existir en los datos. En concreto, el estadístico de contraste que se propone es:

$$q_t(²) = \frac{1}{s_t(²)} \left[\max_{1A¬At} \sum_{a=1}^{¬} (¹_a \quad ¹) \quad \min_{1A¬At} \sum_{a=1}^{¬} (¹_a \quad ¹) \right] \qquad [12]$$

donde ² es un parámetro de truncamiento, ¹ es la media muestral de las T observaciones, y $s_t(²)$ es un estimador de la desviación típica de $¹_µ$ definido como:

$$s_t(²) = \sqrt{¤(0) + 2\sum_{«=1}^{²} ¸_«(²)¤(«)[} \quad \text{con} \quad ¸_«(²) = 1 \quad \frac{«}{² + 1}[\quad \text{con} \quad ² \setminus t$$

donde ¤(0) y ¤(«) denotan la varianza y la autocovarianza muestral de orden « de $¹_µ$, $¤(0) = \frac{1}{t}\sum_{µ=1}^{t}(¹_µ \quad ¹)^2$ y $¤(«) = \frac{1}{t}\sum_{µ=1}^{t P «}(¹_µ \quad ¹)(¹_{µ+«} \quad ¹)$,

respectivamente. El estadístico $q_t(²)$ en [12] se denomina "rango reescalado modificado", y para $² = 0$ se reduce al estadístico $r]s$ de Hurst. Valores extremos de $q_t(²)$, en ambos sentidos, se consideran indicativos de la posible existencia de memoria larga en la serie $¹_µ$. Bajo ciertas condiciones que incluyen una clase muy general de procesos con memoria corta, Lo (1991) deriva la distribución asintótica del estadístico $v_t(²) = \frac{q_t(²)}{\sqrt{t}}$, y proporciona los valores críticos del contraste para los niveles de significación habituales. Además, este autor comprueba que el contraste tiene una potencia razonable, que aumenta con t y disminuye con $²$, frente a alternativas del tipo ARFIMA(0[¥[0). Frente a esta clase de modelos, el contraste detecta mejor valores negativos de ¥ que alternativas del tipo $¥ \wedge 0$. Sin embargo, las simulaciones de Cheung (1993b) sugieren que este contraste es muy sensible a la existencia de correlación bajo la hipótesis nula: por ejemplo, un MA(1) con valores grandes del parámetro sesga el contraste hacia alternativas $¥ \setminus 0$. Por otra parte, Teverovsky, Taqqu y Willinger (1999) comprueban que, para valores grandes de $²$, el estadístico $v_t(²)$ tiende a aceptar la hipótesis nula de memoria corta aún cuando exista memoria larga en los datos.

Robinson (1991) desarrolla un estadístico basado en los multiplicadores de Lagrange (lm) para contrastar la hipótesis de ruido blanco frente a una familia muy general de alternativas que incluye a la clase de los procesos ARFIMA(0[¥[0). En este caso, el estadístico admite una expresión muy sencilla de la forma $#_1 = 5_1^2$, donde $5_1 = \sqrt{\frac{6t}{9^2}} \sum_{«=1}^{t-1} \frac{³(«)}{«}$, siendo $³(«) = ¤(«)]¤(0)$ la autocorrelación muestral de orden $«^2$. Bajo la hipótesis $h_0 : ¥ = 0[5_1$ tiene una distribución asintótica normal $n(0[1)$, y por tanto, $#_1$ se distribuye asintóticamente como una $?_1^2$. Estos dos estadísticos permiten contrastar alternativas unilaterales y bilaterales del tipo $¥ \wedge 0[¥ \setminus 0$ ó $¥ \neq 0$. Cheung (1993b) comprueba que, en muestras finitas, este contraste detecta adecuadamente alternativas fraccionales, aunque es muy sensible a la presencia de correlación serial en los datos: un parámetro AR(1) o MA(1) hace que el contraste rechace h_0 con muchísima más frecuencia que el nivel de significación nominal. Robinson (1994a) sugiere una forma de robustecer el estadístico $#_1$ contra la presencia de autocorrelación débil en la hipótesis nula, y Agiakloglou y Newbold (1994) presentan una

[2] Además de este contraste, Robinson (1991) propone otra variante que es robusta a la heteroscedasticidad condicionada de las perturbaciones.

generalización del mismo para contrastar h_0:ARMA(±[²) frente a la alternativa h_1:ARFIMA(±[¥[²).

Posteriormente, Lobato y Robinson (1998) han propuesto un contraste semiparamétrico para la hipótesis I(0) frente a alternativas de integración fraccional, basado en un estadístico lm aplicado a la función objetivo del método de estimación de Robinson (1995b). Bajo ciertas condiciones se demuestra que, en procesos lineales multivariantes de dimensión ±, el estadístico lm se distribuye, bajo la hipótesis nula ($h_0 : ¥ = 0$), como una distribución $?^2_±$. Además, para el caso univariante (± = 1), el estadístico del contraste puede escribirse como lm $= µ^2$, donde t se define como:

$$µ = \frac{R}{±/®\, T\sum_{«=1}^{®} =_« i(5_«)]\sum_{«=1}^{®} i(5_«)U} \quad \frac{S}{[\,con\ =_« = -^{-}\ «\quad \frac{1}{®}\sum_{«=1}^{®} -^{-}\ «}$$

siendo ® un parámetro de truncamiento, con $0 \setminus ® \setminus t\ ]2$. En este caso, la hipótesis $h_0 : ¥ = 0$ se rechaza en favor de la alternativa de memoria larga ($h_1 : ¥ \wedge 0$), cuando el estadístico µ toma un valor significativamente positivo en la cola derecha de la distribución normal n (0[1). Este contraste parece tener buenas propiedades de eficiencia frente a alternativas fraccionales, comparado con otros contrastes semiparamétricos.

A continuación, se describen brevemente algunos contrastes de raíces unitarias, en los que $h_0 : ¥ = 1$, que pueden enmarcarse en el ámbito de los procesos fraccionalmente integrados. Sowell (1990) deriva la distribución asintótica del estadístico del contraste de Dickey-Fuller para series fraccionalmente integradas i(¥), con $1]2 \leqslant ¥ \setminus 3]2$, y demuestra que ésta depende del orden de integración ¥. Diebold y Rudebush (1991) analizan las propiedades de este contraste en muestras finitas y comprueban que tiene muy poca potencia frente a alternativas fraccionales. En consecuencia, será muy difícil que permita discriminar entre un proceso integrado I(1) y un proceso I(¥) con $¥ \setminus 1$. Hassler y Wolters (1994) realizan un estudio similar para el contraste de Dickey-Fuller aumentado y prueban que éste tiene incluso menos potencia frente a alternativas I(¥). Dolado y Marmol (1997) extienden estos resultados al contraste secuencial de Dickey-Pantula, permitiendo así el análisis de procesos integrados con más de una raíz unitaria.

Lee y Schmidt (1996) consideran el contraste KPSS de Kwiatkowski *et al.* (1992), en el que la hipótesis nula es $h_0 : ¥ = 0$, y prueban

que es consistente frente a procesos con memoria larga estacionarios I(¥), con 1]2 \ ¥ \ 1]2 y ¥ ≠ 0. Lee y Amsler (1997) extienden estos resultados a procesos fraccionalmente integrados no estacionarios (1]2 \ ¥ \ 1) y demuestran que el contraste KPSS no puede distinguir de forma consistente estos procesos de un proceso con raíz unitaria (¥ = 1).

Marmol (1998) revisa los resultados anteriores y prueba que los contrastes unilaterales de Dickey-Fuller (h 0 : ¥ = 1) y KPSS (h 0 : ¥ = 0) son consistentes frente a alternativas fraccionales del tipo h 1 : ¥ \ 1 y h 1 : ¥ ^ 0, respectivamente, pero la utilización combinada de ambos contrastes sólo conlleva a la decisión correcta, para todo ¥ ^ 1]2, si éstos se realizan como contrastes bilaterales.

Robinson (1994a) propone un contraste de los multiplicadores de Lagrange para la hipótesis de raíz unitaria y otras formas de no estacionariedad, frente a alternativas que incluyen a los procesos fraccionalmente integrados. Cuando la perturbación del modelo es un ruido blanco con media cero, este contraste tiene propiedades estándar de eficiencia y tiene una distribución asintótica ?^{2} bajo la hipótesis nula. Robinson (1994a) deriva además la distribución asintótica del estadístico cuando la perturbación del modelo está débilmente correlacionada. En muestras finitas, este contraste compara favorablemente con otros contrastes de raíces unitarias, y permite contrastar no sólo la hipótesis ¥ = 1 sino otros valores del parámetro ¥. Gil-Alana y Robinson (1997) aplican este contraste al conjunto de series macroeconómicas de Nelson y Plosser (1982) y obtienen conclusiones diferentes sobre la existencia de memoria larga, dependiendo del tipo de serie analizada y del modelo utilizado.

Recientemente, Dolado, Gonzalo y Mayoral (1999) han propuesto un contraste para la hipótesis h 0 : i (¥$_0$) frente a h 1 : i (¥$_1$), donde ¥$_1$ \ ¥$_0$, basado en el estadístico de Dickey-Fuller. El estadístico del contraste requiere la estimación del parámetro ¥ bajo la alternativa, y para ello estos autores proponen adicionalmente un estimador de mínima distancia de ¥ con buenas propiedades, tanto asintóticas como en muestras finitas.

Alternativamente, un instrumento muy utilizado para detectar memoria larga en una serie temporal es la función de autocorrelación muestral. Dado que, como vimos anteriormente, las autocorrelaciones de un proceso con memoria larga decaen muy lentamente hacia ce-

ro, dicho comportamiento en las autocorrelaciones muestrales suele considerarse como indicio de integración fraccional. Sin embargo, las autocorrelaciones muestrales son estimadores negativamente sesgados de sus análogos poblacionales en presencia de memoria larga, y esto puede dificultar su interpretación; ver, por ejemplo, Newbold y Agiakloglou (1993), Ray (1993a) y Hosking (1996). Nótese que la presencia de memoria larga afecta incluso a las propiedades de la media muestral. Cuando ¥ \ 0, la velocidad de convergencia de ésta a la media poblacional es mayor que en el caso estacionario, mientras que cuando ¥ ^ 0, dicha convergencia es más lenta; véanse a este respecto, por ejemplo, Beran (1994) y Eisinga, Franses y Ooms (1999). Por lo tanto, puede concluirse que el parámetro ¥ tiene efectos importantes sobre las propiedades asintóticas de los estadísticos de contraste más habituales.

Finalmente, nótese que, como señalan Crato y Ray (1996) y Smith, Taylor y Yadav (1997), el sesgo en la estimación del parámetro ¥ puede afectar también a la identificación del modelo ARFIMA que mejor representa el comportamiento dinámico de una determinada serie, y en consecuencia, los criterios de selección habituales pueden identificar un modelo incorrecto.

5. Predicción de modelos ARFIMA

En esta sección, vamos a considerar el problema de la predicción de valores futuros de una variable $^{1}_{t+\neg}$ generada por un modelo ARFIMA (±[¥[²), de la que se han observado valores hasta el momento t , es decir {$^{1}_{1}$[ZZZ[$^{1}_{t}$}. Si el criterio de predicción es la minimización del error cuadrático medio (ECM), la predicción óptima es la esperanza condicional, dada por

$$\hat{}^{1}_{t+\neg} = [1 \quad b(l)]\,\hat{}^{1}_{t+\neg P 1} \qquad [13]$$

donde $b(l) = 2(l)^{P1}>(l)(1 \quad l)^{¥}$ y $\hat{}^{1}_{t+«} = {}^{1}_{t+«}$ para « 4 0; ver, por ejemplo, Peiris y Perera (1988). El correspondiente ECM de la predicción viene dado por:

$$e(^{1}_{µ+\neg} \quad \hat{}^{1}_{t+\neg})^2 = ;^{2}_{1}\sum_{ª=0}^{\neg P 1})^{2}_{ª} \qquad [14]$$

donde $)_{ª}$ son los coeficientes de la representación MA(Q) de la serie $^{1}_{µ}$.

La principal diferencia entre las predicciones puntuales basadas en modelos estacionarios con memoria corta y las basadas en modelos

con memoria larga es que las primeras convergen exponencialmente a la media muestral mientras que las segundas lo hacen a un ritmo hiperbólico más lento.

Nótese que el predictor definido en [13] para un modelo con memoria larga es óptimo cuando se dispone de observaciones en el pasado hasta el infinito. Sin embargo, en la práctica se dispone de un número finito de observaciones, por lo que hay dos alternativas para obtener predicciones. La primera, propuesta por Beran (1994), se basa en generar predicciones óptimas teniendo en cuenta que el número de observaciones es finito. En este caso, se requiere una expresión de las autocorrelaciones y, por lo tanto, que la serie $^{1}{}_{\mu}$ sea estacionaria. Alternativamente, se puede aproximar el modelo por su representación autorregresiva e igualar a cero todos los valores de $^{1}{}_{\mu}$ previos a la primera observación disponible; ver, por ejemplo, Peiris y Perera (1988) y Ray (1993b). Beran (1994) prueba que, en muestras finitas, las predicciones que proporciona este procedimiento pueden ser muy distintas de las óptimas que tienen en cuenta la dimensión finita de la muestra. Sin embargo, Ray (1993b) calcula el incremento porcentual en el error de predicción cuando un ruido blanco fraccional es modelizado mediante un modelo AR(±) y concluye que dicho modelo no tiene un mal comportamiento. Geweke y Porter-Hudack (1983), Ray (1993a), Sutcli;e (1994), Tiao y Tsay (1994) y Franses y Ooms (1997) tampoco encuentran mejoras sustanciales en la predicción puntual de series reales con memoria larga y sugieren que, en la práctica, los modelos ARMA pueden competir en predicción con los modelos ARFIMA. Crato y Ray (1996), basándose en los resultados de un extenso estudio de simulación, también concluyen que los modelos ARMA habituales proporcionan predicciones razonables. Únicamente cuando el número de observaciones es elevado y la persistencia es muy fuerte, puede estar justificado utilizar modelos ARFIMA para predecir. Finalmente, Eisinga, Franses y Ooms (1999), en una aplicación a la predicción de la orientación política de varios partidos en Holanda, señalan que, cuando el tamaño muestral es suficientemente grande, las diferencias entre ambas predicciones no son importantes, aunque sí hay diferencias sustanciales en las predicciones, dependiendo de que se analicen las variables en niveles o diferenciadas.

Sin embargo, las diferencias fundamentales en la predicción de series con memoria larga están en los intervalos de predicción. Beran (1994) señala que los intervalos de predicción basados en modelos con memoria corta pueden no tener las coberturas deseadas. El parámetro

¥ es fundamental en la determinación de la velocidad de crecimiento del ECM en [14]. Diebold y Lindner (1996) prueban que si el proceso es estacionario (¥ \ 1]2), dicho error converge, cuando ¬ A Q, a la varianza marginal del proceso $^{1}{}_{\mu}$, es decir, tiene el mismo comportamiento que en series estacionarias con memoria corta. Sin embargo, el caso no estacionario es más interesante. En concreto, si 1]2 \ ¥ \ 3]2, el ECM se incrementa con el horizonte de predicción, ¬, a una tasa ¤¬$^{2¥P1}$, donde ¤ es una constante positiva. Esto significa que, en el modelo no estacionario con reversión a la media (1]2 \ ¥ \ 1), la tasa de incremento del ECM es intermedia entre el caso estacionario y el de series I(1), donde la tasa convergencia es de orden ¬. Por lo tanto, si 1]2 \ ¥ \ 1, los intervalos de predicción basados en modelos ARIMA serían, o bien asintóticamente muy estrechos si se eligiera ¥ = 0 o innecesariamente anchos si ¥ = 1. En este sentido, Crato y Taylor (1996) presentan resultados sobre los ECM que se obtienen cuando una serie generada como un ruido blanco fraccional, es decir, ARFIMA(0[¥[0), es erróneamente diferenciada. En particular, si ¥ R (1]2[3]2), el predictor basado en el paseo aleatorio tiene ECM muy superiores especialmente cuando ¥ ∧ 1. Si ¥ R (1]2[1]2), el incremento relativo en el error de predicción es mayor cuanto menor sea el horizonte de predicción y converge a uno cuando dicho horizonte se incrementa. En este último caso, si se predice mediante un paseo aleatorio con deriva, los efectos son todavía mas graves, en el sentido de que el ECM se incrementa con $¬^2$. Diebold y Lindner (1996) comparan empíricamente los intervalos de predicción obtenidos mediante un modelo de memoria larga ajustado a una serie mensual de tipos de interés, con los que se obtienen mediante un paseo aleatorio. Aunque las predicciones obtenidas mediante ambos modelos no son muy diferentes, los intervalos correspondientes al paseo aleatorio son innecesariamente anchos en relación a los obtenidos con el modelo con integración fraccional. Beran y Ocker (1999) y Brodsky y Hurvich (1999) llegan a la misma conclusión analizando series reales de tipos de cambio diarios y series simuladas, respectivamente. Finalmente, Reisen y Lopes (1999) presentan algunas simulaciones sobre el comportamiento de los intervalos de predicción en modelos ARFIMA(±[¥[^{2}).

6. Modelos con memoria larga en la varianza condicionada

Muchas series temporales financieras observadas con una periodicidad muy alta suelen caracterizarse por tener escasa o nula estructura dinámica en la media, exceso de curtosis y correlaciones significativas en

la serie de los cuadrados. Estas características se han relacionado habitualmente con cambios temporales en la varianza condicionada. Entre los modelos propuestos para representar la evolución dinámica de la varianza condicionada, se distinguen dos grandes grupos: los modelos tipo ARCH (AutoRegressive Conditional Heteroscedasticity), basados en la propuesta inicial de Engle (1982), y los modelos de volatilidad estocástica SV (Stochastic Volatility) propuestos por Taylor (1986). En ambos casos, la serie de interés, $\{^{o}{}_{\mu}\}_{\mu=1}^{t}$, se modeliza a través de la ecuación:

$$^{o}{}_{\mu} = ;{}_{\mu}B_{\mu} \qquad [15]$$

donde B_{μ} es un ruido blanco independiente e igualmente distribuido (I.I.D.) con media cero y varianza unitaria, y $;{}_{\mu}$ es un factor positivo denominado volatilidad. En los modelos ARCH, $;{}_{\mu}$ es función de las observaciones pasadas de la serie, mientras que en los modelos SV, $;{}_{\mu}$ es un proceso estocástico no observable.

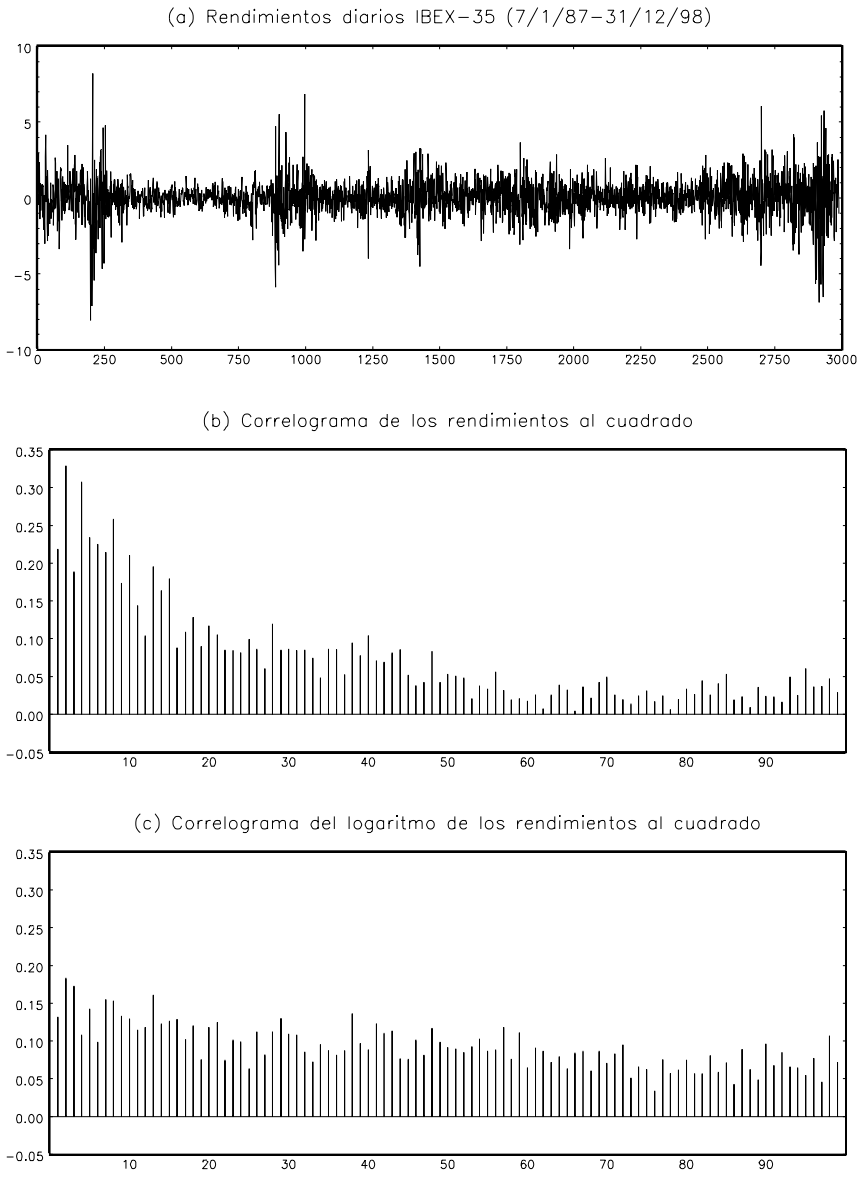
En las aplicaciones empíricas de estos modelos a series financieras observadas con gran frecuencia, es casi sistemático estimar un fuerte grado de persistencia en la volatilidad. Esta regularidad empírica motivó la introducción de modelos con raíces unitarias en la volatilidad. Estos modelos se caracterizan, entre otras cosas, porque el efecto de un *shock* en la varianza condicionada persiste para siempre e influye en las predicciones futuras de la volatilidad para todos los horizontes de predicción. Sin embargo, algunos trabajos recientes cuestionan la validez de los modelos integrados, precisamente porque el grado extremo de persistencia que inducen en la volatilidad no refleja las características observadas empíricamente en ciertas series financieras, que sugieren más bien la existencia de memoria larga en la volatilidad. En este sentido, Ding, Granger y Engle (1993) analizan una serie diaria del índice bursátil S&P500 y observan que las autocorrelaciones de ciertas transformaciones de los rendimientos absolutos tienden a persistir, incluso para retardos muy altos, pero finalmente convergen hacia cero. Lobato y Savin (1998) aplican un contraste semiparamétrico de memoria larga a la misma serie y encuentran también una fuerte evidencia de memoria larga en la serie de los cuadrados, aunque no en la propia serie de rendimientos. Dacorogna *et al.* (1993) observan un comportamiento similar en una serie de datos del tipo de cambio dólar frente al marco alemán, observados cada veinte minutos. Estos autores ajustan dos funciones, una hiperbólica y otra exponencial, al correlograma de los rendimientos absolutos y comprueban que el decrecimiento hi-

perbólico, típico de los procesos fraccionalmente integrados, se ajusta mejor que el exponencial. Resultados similares aparecen en Andersen y Bollerslev (1997). Crato y de Lima (1994) contrastan la persistencia en la volatilidad de varias series de índices bursátiles, aplicando los contrastes de memoria larga GPH y r]s a la serie de los cuadrados de los rendimientos y a la serie de sus logaritmos, y concluyen que, efectivamente, existe dependencia a largo plazo en muchas de ellas. Estos resultados se confirman posteriormente en Breidt, Crato y de Lima (1998). Olmeda (1998) aplica estos mismos contrastes a series del mercado bursátil español y concluye que hay cierta evidencia de memoria larga en las volatilidades.

Para ilustrar este comportamiento de la volatilidad de las series financieras, el Gráfico 2(a) representa la evolución de una serie de rendimientos diarios del índice IBEX35 de la Bolsa de Madrid, observados entre 7/1/1987 y 30/12/1998, con un total de 2991 observaciones. La serie se ha filtrado previamente de atípicos y de cualquier estructura dinámica de la media. Los Gráficos 2(b)-2(c) representan los correlogramas de los cuadrados de los rendimientos y del logaritmo de los cuadrados, respectivamente. Este gráfico muestra la existencia de una fuerte dependencia en los momentos de segundo orden, con correlaciones positivas en la serie de los cuadrados, incluso en retardos muy alejados. Esta persistencia se agudiza en la serie del logaritmo de los cuadrados, que presenta correlaciones más pequeñas en magnitud pero con un decrecimiento muy lento hacia cero típico de los procesos con memoria larga. El Cuadro 3 recoge los resultados de algunos de los contrastes de memoria larga descritos en la sección 4, aplicados a las series $\circ^2_\mu$ y $-^{-}(\circ^2_\mu)$. A juzgar por estos resultados, el rechazo de la hipótesis de memoria corta parece incuestionable en ambas series, aunque la evidencia de memoria larga es más clara en la serie $-^{-}(\circ^2_\mu)$ que en $\circ^2_\mu$.

Entre los distintos modelos propuestos para representar la posible existencia de memoria larga en los momentos condicionados de segundo orden, en esta sección se describen aquellos que han tenido mayor relevancia en aplicaciones empíricas. Para cada uno de ellos, se describen sus características más importantes y se discuten sus posibles ventajas e inconvenientes. Finalmente, se revisan brevemente otros modelos con memoria larga en la varianza condicionada.

GRÁFICO 2
Serie de rendimientos diarios del IBEX-35 (7/1/1987-30/12/1998) junto con los correlogramas de los cuadros y del logaritmo de los cuadrados de dicha serie.



CUADRO 3
Contrastes unilaterales de memoria larga (H_0:d=0 frente H_1:d>0) para la volatilidad del IBEX-35.

Contraste GPH (1983)								
Serie	y_t^2				$ln(y_t^2)$			
$m=$	$T^{0.45}$	$T^{0.5}$	$T^{0.6}$	$T^{0.5}$	$T^{0.45}$	$T^{0.5}$	$T^{0.6}$	$T^{0.5}$
$\ell=$				$T^{0.1}$				$T^{0.1}$
$\hat{d}$ GPH	0.2818	0.3802	0.3992	0.3725	0.6241	0.5466	0.4574	0.5546
t	2.8615	4.5522	7.7639	3.5086	6.1070	6.0224	6.5213	4.8630
p-valor	0.0021	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000
Contraste R/S de Lo (1991)								
Serie	y_t^2				$ln(y_t^2)$			
$q=$	$q=25$	$q=50$	$q=100$	$q=q^*=9$	$q=25$	$q=50$	$q=100$	$q=q^*=9$
$V_T(q)$	2.2533	1.8780	1.6247	3.0800	2.5711	2.0263	1.5795	3.9117
p-valor	0.0015	0.0227	0.0974	0.0000	0.0001	0.0084	0.1222	0.0000
Contraste LM de Robinson y Lobato (1998)								
Serie	y_t^2				$ln(y_t^2)$			
$m=$	$m=30$	$m=50$	$m=70$	$m=100$	$m=30$	$m=50$	$m=70$	$m=100$
t	2.6764	4.8610	6.2038	8.6390	5.6971	9.0527	10.747	14.035
p-valor	0.0037	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

6.1. Modelos GARCH fraccionalmente integrados

Robinson (1991) sugirió la primera extensión del modelo GARCH de Bollerslev (1986) capaz de producir memoria larga en los cuadrados de la serie, pero son Baillie, Bollerslev y Mikkelsen (1996) quienes desarrollan el modelo y lo denominan GARCH fraccionalmente integrado (FIGARCH). En el modelo GARCH(±[²), la varianza condicional de la serie, ;$^2_{\mu}$, es una función lineal de su propio pasado y de las observaciones pasadas de la serie °$_{\mu}$, de la siguiente forma:

$$;^2_{\mu} = +_0 + +(1)\ {}^{o}{}^2_{\mu} + ,(1);^2_{\mu} \qquad [16]$$

donde $+(1) = \sum_{a=1}^{2} +_a 1^a[$, $(1) = \sum_{«=1}^{\pm} ,_« 1^«[$ ² ^ 0[± 5 0[$+_0$ ^ 0[$+_a$ 5 0[ª = 1[ZZ[²[, $_«$ 5 0[« = 1[ZZZ[±[3]. El modelo GARCH definido por

[3] Aunque estas restricciones se establecen para garantizar que la varianza sea positiva, Nelson y Cao (1992) demuestran posteriormente que la positividad de la varianza está asegurada bajo condiciones más débiles.

las ecuaciones [15] y [16] puede interpretarse como un proceso ARMA para la serie de los cuadrados, de la forma:

$$[1 \quad +(l) \quad ,(l)]\,{}^{o}{}^{2}_{\mu} = +_0 + [1 \quad ,(l)]\,=_{\mu} \qquad [17]$$

donde $=_{\mu}$ es un ruido blanco definido como $=_{\mu} = {}^{o}{}^{2}_{\mu} \quad ;^{2}_{\mu} = ;^{2}_{\mu}(B^{2}_{\mu} \quad 1)$.

El modelo FIGARCH no es más que una extensión del modelo anterior, que incluye el operador de diferencias fraccionales, $(1 \quad l)^{¥}$, en la parte autorregresiva de la ecuación [17], de tal forma que el modelo resultante puede interpretarse como un proceso ARFIMA para la serie de los cuadrados, de la forma:

$$>(l)(1 \quad l)^{¥}\,{}^{o}{}^{2}_{\mu} = +_0 + [1 \quad ,(l)]\,=_{\mu} \qquad [18]$$

donde ¥ es un número real 0 4 ¥ 4 1 y todas las raíces de los polinomios $>(l)$ y $[1 \quad ,(l)]$ están fuera del círculo unidad. Alternativamente, la ecuación de la varianza condicionada en el modelo FIGARCH puede escribirse como:

$$\begin{aligned} ;^{2}_{\mu} \quad &= \quad [1 \quad ,(1)]^{P1}\,+_0 + \left\{1 \quad [1 \quad ,(l)]^{P1}\,>(l)(1 \quad l)^{¥}\right\}\,{}^{o}{}^{2}_{\mu} = \\ &\qquad [1 \quad ,(1)]^{P1}\,+_0 + 5(l)\,{}^{o}{}^{2}_{\mu} \end{aligned} \qquad [19]$$

Al permitir órdenes de integración fraccional, el modelo FIGARCH proporciona una gran flexibilidad en la modelización de la dependencia temporal de la varianza condicionada, e incluye como casos particulares el modelo GARCH (¥ = 0) y el modelo GARCH integrado o IGARCH (¥ = 1). El modelo es, por tanto, el análogo para la varianza condicionada de los modelos ARFIMA para la media revisados en la sección 2.

Para que el modelo FIGARCH esté bien definido y se garantice la no negatividad de la varianza, los parámetros del modelo deben cumplir una serie de restricciones que, salvo en casos muy concretos, no son fáciles de establecer y suponen además una gran desventaja en la estimación del modelo. Respecto a las condiciones de estacionariedad, el modelo FIGARCH sólo es débilmente estacionario cuando ¥ = 0, en cuyo caso se reduce al GARCH estándar. Sin embargo, Baillie, Bollerslev y Mikkelsen (1996) argumentan que el modelo es estrictamente estacionario y ergódico, de la misma forma que lo es el IGARCH.

Para ilustrar el comportamiento de las series generadas por modelos FIGARCH y de las series de sus cuadrados, el Gráfico 3 representa

sendas realizaciones de un proceso GARCH(1,1) con $\{+_0 = 0.1[\ +_1 = 0.125[\ ,_1 = 0.85\}$, dos procesos FIGARCH(1[¥[0) con parámetros $\{+_0 = 0.1[\ ¥ = 0.4[\ ,_1 = 0.25\}$ y $\{+_0 = 0.1[\ ¥ = 0.75[\ ,_1 = 0.7\}$, respectivamente, y un proceso IGARCH ($¥ = 1$) con $\{+_0 = 0.1[\ ,_1 = 0.85\}$. Estos valores de los parámetros son los que utilizan Baillie, Bollerslev y Mikkelsen (1996) y Bollerslev y Mikkelsen (1996) en sus simulaciones. Las series FIGARCH simuladas se han obtenido truncando el polinomio $5(l)$ de la ecuación [19] en el retardo $j = 3000$ y eliminando las primeras 4000 observaciones para evitar problemas de valores iniciales. Cada una de las series representadas se acompaña de la función autocorrelación muestral de la serie de sus cuadrados. Para el modelo GARCH(1,1), el único estacionario, se representa además la función de autocorrelación poblacional de $°^2_{¼}$, cuya expresión aparece en Bollerslev (1988). El Gráfico 3 muestra claramente la dependencia a largo plazo en la serie de los cuadrados para los modelos FIGARCH, frente al decrecimiento exponencial del modelo GARCH ($¥ = 0$). Sin embargo, no parece que vaya a ser tarea sencilla poder discriminar entre modelos FIGARCH con valores $¥ \setminus 1]2$ y $¥ \wedge 1]2$. Recuérdese que el proceso FIGARCH no es estacionario ni siquiera en el intervalo $¥ \setminus 1]2$.

Una de las ventajas del modelo FIGARCH es la flexibilidad que proporciona para estudiar el grado de persistencia de los *shocks* sobre la volatilidad. Al igual que en los modelos para la media, Baillie, Bollerslev y Mikkelsen (1996) proponen cuantificar este grado de persistencia utilizando la función de respuesta al impulso acumulativa. En los modelos FIGARCH, los valores de esta función serán los coeficientes del polinomio $¤(l) = 1 + ¤_1 l + ¤_2 l^2 + \mathbb{ZZ}$ tal que $°^2_{¼} = 0 + ¤(l)=_{¼}$, y el coeficiente $¤_?$ medirá el impacto a largo plazo que tiene una innovación unitaria sobre el nivel de la serie de los cuadrados. Utilizando la expresión [18] del modelo FIGARCH, Baillie, Bollerslev y Mikkelsen (1996) prueban que $¤_? = f(¥ \quad 1[\,1[\,1;1)>(1)^{P1}[1 \quad ,(1)]$, y en virtud de los resultados de Gradshteyn y Ryzhik (1980, p. 1040) sobre la función f, resulta que:

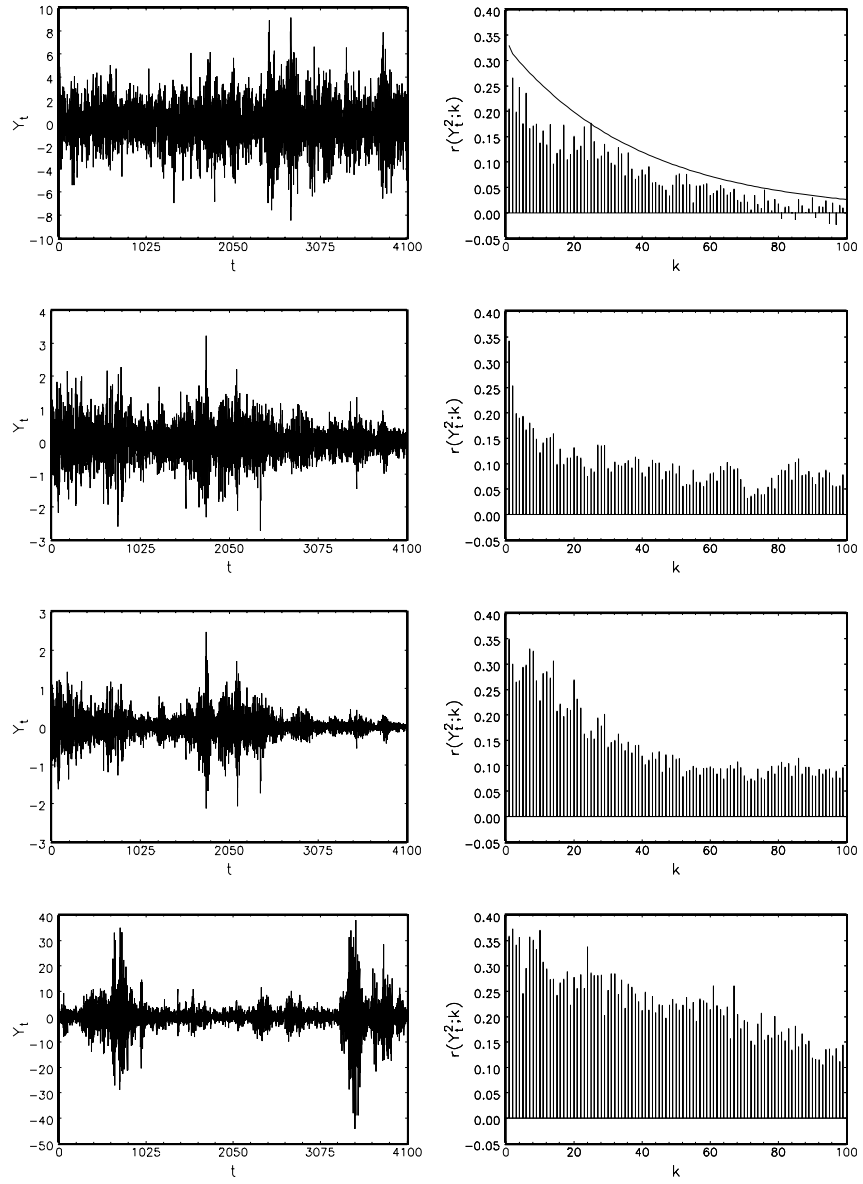
$$¤_? = \begin{cases} 0 & ´^a \quad 0 \ 4 \ ¥ \setminus 1 \\ >(1)^{P1}[1 \quad ,(1)] & ´^a \quad ¥ = 1 \end{cases}$$

Por tanto, en el modelo IGARCH ($¥ = 1$), el efecto de los *shocks* sobre la volatilidad persiste indefinidamente ($¤_? \neq 0$), mientras que en los modelos GARCH ($¥ = 0$) y FIGARCH ($0 \setminus ¥ \setminus 1$), el efecto del ´©°¤¬ finalmente desaparece ($¤_? = 0$). Sin embargo, existe una diferencia sustancial en la velocidad con la que dicho efecto desaparece en ambos

GRÁFICO 3
Series y_t generadas por diferentes modelos FIGARCH junto con la función de autocorrelación muestral de la serie de sus cuadros. De arriba abajo, los modelos simulados son GARCH(1,1) con

$\{+_0 = 0[1[\ +_1 = 0[125[\ ,_1 = 0[85\}[\,\mathrm{FIGARCH}(1[\,¥[\,0)\,\mathrm{con}\,\{+_0 = 0[1[\,¥ = 0[4[\ ,_1 = 0[25\}$

$\mathrm{y}\,\{+_0 = 0[1[\,¥ = 0[75[\ ,_1 = 0[7\}[\,\mathrm{el\,GARCH\,con}\,\{+_0 = 0[1[\ ,_1 = 0[85\}ⵈ$



modelos. Por ejemplo, en un modelo GARCH(1,1) estacionario ($+_1 + ,_1 \setminus 1$), se demuestra que los valores de la función respuesta al impulso acumulativa son de la forma:

$$¤_¬ = +_1(+_1 + ,_1)^{¬P1} \qquad [20]$$

para $¬ \; 5 \; 1$ y $¤_0 = 1$, mientras que en un modelo FIGARCH(1[¥[0), son de la forma:

$$¤_¬ = \left[(1 \quad ,_1) \quad \frac{(1 \quad ¥)}{¬}\right] \frac{(¬ \quad 1 + ¥)}{(¬) \;\; (¥)} \; \underline{8} \; \frac{(1 \quad ,_1)}{(¥)} ¬^{¥P1}[\,'^a \; ¬ \; A \;\; Q \qquad [21]$$

para $¬ \; 5 \; 1$ y $¤_0 = 1$. De las expresiones [20] y [21] se deduce que, mientras en un modelo GARCH(1,1), los coeficientes $¤_¬$ decrecen exponencialmente hacia cero, en un modelo FIGARCH(1[¥[0), el efecto de un *shock* desaparece más despacio, a un ritmo hiperbólico determinado por el valor del parámetro ¥. En consecuencia, el parámetro ¥ se identifica con el ritmo al que se propagan los *shocks* en la volatilidad y no con el impacto último que éstos tienen sobre la predicción de la varianza condicionada, que será cero tanto en los modelos GARCH estacionarios ($¥ = 0$) como en los FIGARCH ($0 \setminus ¥ \setminus 1$).

Para estimar el modelo FIGARCH, Baillie, Bollerslev y Mikkelsen (1996) proponen el método de máxima verosimilitud en el dominio temporal, bajo la hipótesis de normalidad, y comprueban que los estimadores tienen buenas propiedades en muestras finitas. Sin embargo, la aplicación de este método tiene varios inconvenientes. En primer lugar, para calcular la función de verosimilitud se requiere truncar el polinomio $(1 \quad l)^{¥}$ en un retardo suficientemente grande que permita capturar las propiedades de memoria larga de la serie. Además, no se conocen aún las propiedades asintóticas de los estimadores.

Alternativamente, Bollerslev y Mikkelsen (1996) proponen una extensión del modelo EGARCH de Nelson (1991) capaz de representar memoria larga en la volatilidad. Esta propuesta, que ya había sugerido el propio Nelson (1991), consiste en modelizar la serie $-^{¯}(;^2_{µ})$ mediante una estructura de retardos en $B_µ$ y en su propio pasado, incluyendo en la parte autorregresiva del modelo, el operador en diferencias fraccionales, $(1 \quad l)^{¥}$. Este modelo, denominado FIEGARCH, tiene ciertas ventajas sobre el modelo FIGARCH, ya que garantiza la no negatividad de la varianza condicionada modelizando su logaritmo, no impone restricciones en los parámetros y permite modelizar respuestas asimétricas de la volatilidad a cambios positivos y negativos en los valores de la

serie. Además, si 0 \ ¥ \ 1]2, el proceso -¯(; $^2_\mu$) es ergódico, estricta y débilmente estacionario, e invertible. A pesar de estas ventajas, el modelo FIEGARCH adolece de los mismos inconvenientes que el modelo FIGARCH con respecto a la estimación. Nótese que ni siquiera para el modelo EGARCH estándar (¥ = 0) existen resultados sobre las propiedades asintóticas de los estimadores propuestos. Bollerslev y Mikkelsen (1999) presentan evidencia de integración fraccional en la volatilidad de observaciones diarias del índice S&P500 y señalan la importancia económica de incorporar la memoria larga en las predicciones. Comparando los precios reales de opciones sobre el índice con precios simulados, estos autores observan que los precios simulados por el modelo FIEGARCH son más cercanos a los reales que los simulados por modelos EGARCH o IEGARCH.

La predicción de valores futuros de series temporales generadas por modelos FIGARCH no ha sido considerada hasta el momento en la literatura, al menos hasta donde nosotras hemos sido capaces de investigar.

6.2. El modelo SV con memoria larga

Como alternativa a los modelos anteriores, Harvey (1998) y Breidt, Crato y de Lima (1998) proponen modelizar la memoria larga en los momentos de segundo orden, utilizando los modelos de volatilidad estocástica SV. En estos modelos, la volatilidad se considera una variable no observable cuyo logaritmo, $\mathbb{C}_\mu$ = -¯(; $^2_\mu$), es un proceso estocástico lineal. En el modelo SV con memoria larga, denominado LMSV, $\mathbb{C}_\mu$ es un proceso ARFIMA. En el caso más sencillo, $\mathbb{C}_\mu$ es un ARFIMA(0[¥[0), y el modelo viene dado por:

$$^{o}{}_{\mu} = ;_{*};_{\mu}\mathrm{B}_{\mu} \qquad [22a]$$

$$(1 \quad l)^{¥}\mathbb{C}_{\mu} = 1_{\mu}[\qquad [22b]$$

donde ;$_*$ es un factor de escala que evita introducir una constante en la ecuación de $\mathbb{C}_\mu$, ¥ es un parámetro que puede tomar cualquier valor en el intervalo 0 4 ¥ 4 1[4] y 1_μ es n i d (0[; $^2_\mu$) e independiente de B_μ. Aunque el supuesto de Normalidad para 1_μ pudiera parecer arbitrario, Andersen *et al.* (2001) sugieren que, empíricamente, la distribución

[4] Conservamos aquí la notación original de Harvey (1998) que considera el modelo LMSV definido en 0 A ¥ A 1, a diferencia de Breidt, Crato y de Lima (1998) que consideran el rango P1]2 \ ¥ \ 1]2, donde el modelo es estacionario e invertible.

Normal es una buena aproximación a la distribución del logaritmo de la volatilidad.

Harvey (1998) deriva las propiedades dinámicas del modelo LMSV en [22]. El proceso $°_µ$ es una martingala en diferencias, y por tanto, $e(°_µ) = 0$ y $Cov(°_µ[°_·) = 0$ para $µ \neq ´$. Además, $°_µ$ es estacionario siempre que lo sea $©_µ$, es decir, cuando ¥ \ 1]2. Con respecto al comportamiento de la serie $°^2_µ$, Breidt, Crato y de Lima (1998) demuestran que, en el intervalo 1]2 \ ¥ \ 1]2, dicha serie es débil y estrictamente estacionaria. Harvey (1998) comprueba que, cuando $;^2_©$ es pequeño y/o $:_©(¬)$ toma valores próximos a uno, la función de autocorrelación de $°^2_µ$ puede aproximarse mediante la siguiente expresión:

$$:_{°^2_µ}(¬) \underline{8} :_©(¬)\frac{exp\left(;^2_©\right) \quad 1}{3\,exp\left(;^2_©\right) \quad 1}[\text{ para } ¬ \; 5 \; 1$$

donde $:_©(¬)$ denota la función de autocorrelación de $©_µ$ y $;^2_©$ su varianza. Por lo tanto, la función de autocorrelación de $°^2_µ$ decrecerá hiperbólicamente hacia cero, al igual que la función de autocorrelación de $©_µ$, pero tomará, en general, valores más pequeños que ésta.

Las propiedades dinámicas del modelo LMSV aparecen de forma más clara en la serie $-^{-}(°^2_µ)$. Considérese la siguiente representación lineal del modelo LMSV:

$$¹_µ = -^{-}(°^2_µ) = 6 + ©_µ + 8_µ \qquad [23a]$$

$$(1 \quad l)^{¥}©_µ = 1_µ \qquad [23b]$$

donde $6 = -^{-}(;^2_t) + e[-^{-}(B^2_µ)]$ y $8_µ = -^{-}(B^2_µ) \quad e[-^{-}(B^2_µ)]$ es un ruido blanco no gaussiano de media 0 y varianza $;^2_8$, cuyas propiedades dependen de las de $B_µ$. Por ejemplo, si $B_µ$ es $n(0[1)[-^{-}(B^2_µ)$ se distribuye como el logaritmo de una $?^2_1$ con media $e[-^{-}(B^2_µ)] = \quad 1[27$ y varianza $;^2_8 = 9^2]2 \underline{8} 4.93$; véase Abramowitz y Stegun (1970, p. 943). La ecuación [23a] representa $¹_µ$ como la suma de dos procesos mutuamente independientes, un proceso lineal ARFIMA(0[¥[0) y un ruido blanco $8_µ$ no gaussiano. Si ¥ \ 1]2, el proceso $©_µ$ es estacionario y se puede calcular fácilmente la varianza, la función de autocovarianzas y la función de autocorrelación de $¹_µ = -^{-}(°^2_µ)$, a partir de las correspondientes funciones de $©_µ$ y $8_µ$. En concreto, se obtienen las siguientes expresiones:

$$-_x(0) = Var(¹_µ) = ;^2_© + ;^2_8 = ;^2_1\frac{(1 \quad 2¥)}{[\ (1 \quad ¥)]^2} + ;^2_8 \qquad [24]$$

$$-_{x}(\neg) = -_{\copyright}(\neg) = ;_{1}^{2}\frac{(1 \quad 2¥)\ (\neg + ¥)}{(¥)\ (1 \quad ¥)\ (\neg + 1 \quad ¥)}[\neg\ 5\ 1 \qquad [25]$$

$$:_{x}(\neg) = \frac{-_{x}(\neg)}{-_{x}(0)} = :_{\copyright}(\neg)\left\{1 + \frac{;_{8}^{2}[\ (1 \quad ¥)]^{2}}{;_{8}^{2}\ (1 \quad 2¥)}\right\}^{P1} [\neg\ 5\ 1 \qquad [26]$$

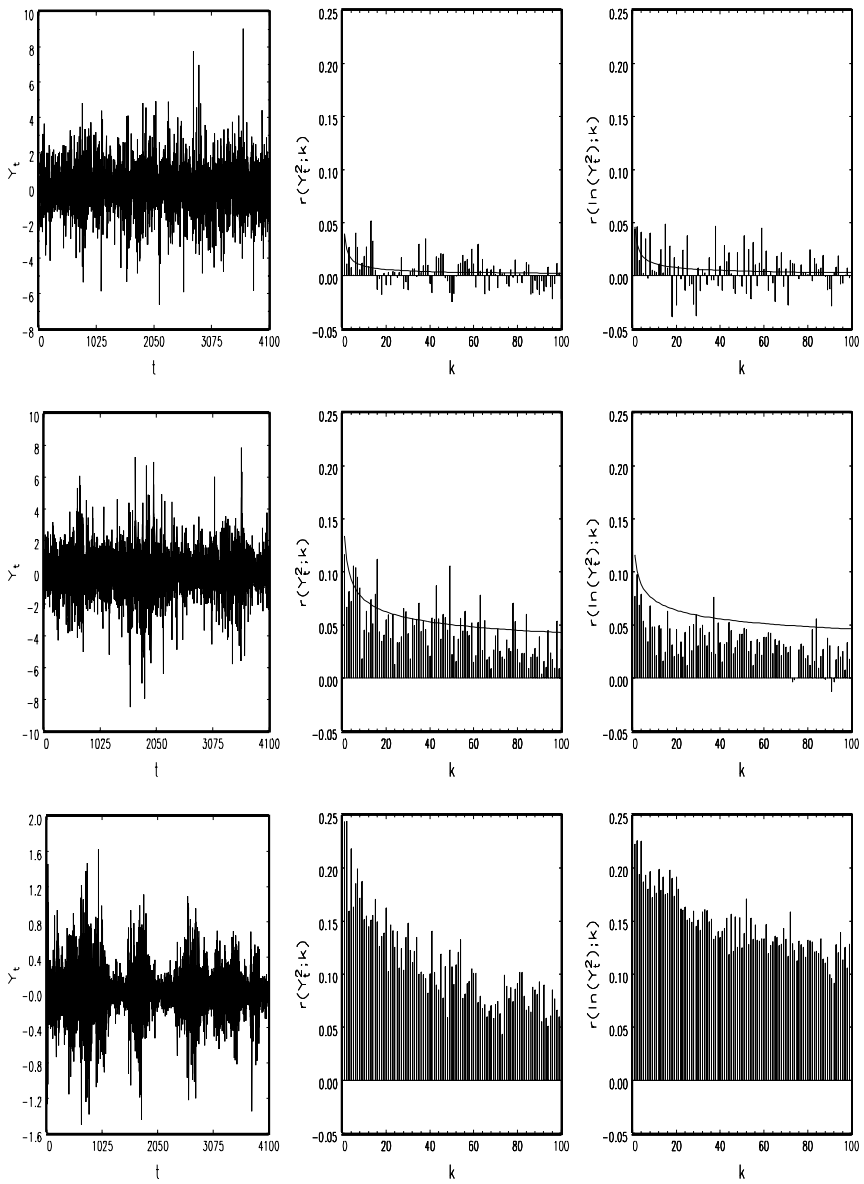
De estas ecuaciones se derivan varias conclusiones importantes. En primer lugar, la ecuación [25] nos dice que las propiedades de memoria que se manifiestan en la función de autocovarianza de $^{1}{}_{\mu}$ son las mismas que las de $\copyright_{\mu}$, y por tanto, si $0 \setminus ¥ \setminus 1]2$, la serie $^{1}{}_{\mu}$, al igual que $\copyright_{\mu}$, tiene memoria larga. Sin embargo, según la ecuación [24], la varianza de $^{1}{}_{\mu}$ es mucho mayor que la de $\copyright_{\mu}$ debido al efecto adicional de la varianza del ruido 8_{μ}. En la ecuación [26] observamos que la función de autocorrelación de $^{1}{}_{\mu}$ es proporcional a la función de autocorrelación de $\copyright_{\mu}$, y por tanto, decrecerá a un ritmo hiperbólico muy lento, aunque tomará siempre valores más pequeños. Por ejemplo, para un ¥ fijo, valores grandes del cociente $;_{8}^{2}];_{1}^{2}$ pueden producir valores tan pequeños de la función de autocorrelación de $^{1}{}_{\mu}$, que resulte realmente difícil distinguir esta serie de un ruido blanco.

El Gráfico 4 representa realizaciones de series generadas por procesos LMSV con parámetros {¥=0.2[$;_{1}^{2}$=1}, {¥=0.4[$;_{1}^{2}$=0.5} y {¥=0.75[$;_{1}^{2}$ =0.1}. Cada una de las series se acompaña de la función de autocorrelación muestral de la serie de los cuadrados y del logaritmo de los cuadrados. En los modelos estacionarios (¥ \ 1]2) se incluye además la función de autocorrelación poblacional de dichas series. En primer lugar, observamos que el comportamiento de las correlaciones diferencia claramente los dos modelos estacionarios del modelo no estacionario (¥=0.75). Además, para el modelo estacionario más persistente (¥=0.4), este gráfico evidencia la dependencia a largo plazo existente en la serie $^{\circ 2}_{\mu}$, y de forma aún más clara en la serie $-^{-}(^{\circ 2}_{\mu})$. Sin embargo, en el modelo con {¥=0.2[$;_{1}^{2} = 1$}, las correlaciones no muestran ese grado de persistencia. De hecho, para valores de $;_{1}^{2}$ más pequeños, apenas podría distinguirse esta serie de un ruido blanco.

La representación lineal del modelo LMSV en [23] facilita asimismo el cálculo de la densidad espectral de la serie $^{1}{}_{\mu}$, cuyo comportamiento también permite caracterizar la propiedad de memoria larga del modelo. En concreto, por la independencia de 1_{μ} y B_{μ}, de la ecuación [23a] se deduce que la densidad espectral de $^{1}{}_{\mu}$ viene dada por la siguiente expresión:

$$§_{x}(5) = \frac{1}{29}\left\{;_{1}^{2}[2(1 \quad ¤^{\circ}{}'5)\ ]^{P¥} + ;_{8}^{2}\right\}$$

GRÁFICO 4
Series y_t generadas por diferentes modelos LMSV junto con la función de autocorrelación, muestral y poblacional, de las series y_t^2 y $ln(y_t^2)$.
De arriba abajo, los modelos simulados son
$\{¥ = 0[2[\ ;\ ^2_1 = 1\}[\ \{¥ = 0[4[\ ;\ ^2_1 = 0[5\}[\ \{¥ = 0[75[\ ;\ ^2_1 = 0[1\}$



Para valores positivos de ¥ esta función no está acotada en las frecuencias bajas, lo que confirma que en el intervalo 0 \ ¥ \ 1]2 la serie $^{1}{}_{\mu}$ tiene la propiedad de memoria larga.

El modelo LMSV básico definido en [22] puede generalizarse de forma que el proceso $©_{\mu}$ sea un proceso ARFIMA(±[¥[^{2}) con componentes autorregresivos y de medias móviles. Las condiciones de estacionariedad y la estructura dinámica de este modelo se obtienen de forma similar a las del modelo LMSV básico, aunque las expresiones de la varianza y la función de autocorrelación de las series $^{o}{}^{2}_{\mu}$ y $\text{-}^{-}(^{o}{}^{2}_{\mu})$ resultan más complicadas; ver Breidt, Crato y de Lima (1998) sobre las propiedades de este modelo.

La estimación del modelo LMSV presenta ciertas dificultades debido a que la distribución condicionada de $^{o}{}_{\mu}$ no tiene una expresión explícita, y por tanto, no es inmediato construir la función de verosimilitud exacta. Harvey (1998) y Breidt, Crato y de Lima (1998) proponen estimar el modelo LMSV maximizando la aproximación discreta de Whittle a la función de verosimilitud gaussiana en el dominio de las frecuencias de la serie $^{1}{}_{\mu} = \text{-}^{-}(^{o}{}^{2}_{\mu})$. Breidt, Crato y de Lima (1998) demuestran que los estimadores resultantes son consistentes, aunque su distribución asintótica no es todavía conocida. Pérez y Ruiz (2001) analizan las propiedades en tamaños de muestra finitos de dichos estimadores y señalan que cuando se estiman modelos con parámetros próximos a las fronteras de la no estacionariedad (¥ 9 1]2) y/o de la homocedasticidad ($;^{2}_{1}$ 9 0), que son los más habituales en aplicaciones con series reales, los estimadores son muy sesgados y tienen mucha dispersión. En estos casos, se necesitan tamaños muestrales extremadamente grandes para que las inferencias sean fiables.

Harvey (1998) propone también un algoritmo que permite obtener la estimación suavizada de la volatilidad en modelos LMSV. Este algoritmo requiere invertir una matriz t ^{1} t , lo cual puede resultar inviable en muestras grandes. En estos casos, Harvey sugiere realizar el suavizado con ponderaciones obtenidas de una muestra más pequeña, y argumenta que con esta estrategia no se pierde mucha precisión. Pérez (2000) aplica esta solución a series simuladas y series reales de tamaños relativamente grandes y comprueba que, efectivamente, proporciona buenas estimaciones de la volatilidad.

Como ilustración, el Cuadro 4 recoge los resultados de la estimación del modelo LMSV en [22] para una serie de rendimientos diarios del

IBEX-35 observados entre el 7/1/1987 y el 30/12/1998, junto con la estimación de un modelo LMSV que incluye además un componente AR(1) en la ecuación de la volatilidad. En ambos casos, el modelo estimado es no estacionario y $\hat{\sigma}_\eta^2$ es pequeño, indicando una evolución suave de los cambios temporales en la volatilidad. El Cuadro 4 recoge también valores del estadístico de Box-Ljung para contrastar incorrelación en la serie de los cuadrados de los residuos estandarizados. Ninguno de estos valores es significativo, por lo que ambos modelos parecen capaces de representar adecuadamente la evolución dinámica de la volatilidad. El Gráfico 5 representa los rendimientos absolutos del índice junto con la estimación suavizada de la volatilidad para el modelo LMSV básico estimado previamente.

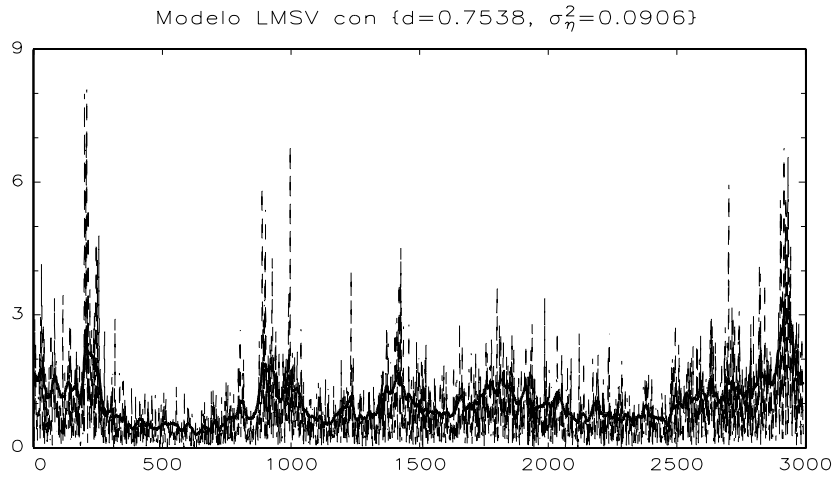
CUADRO 4
Estimación de modelos LMSV para los rendimientos del índice IBEX-35

Modelo LMSV	$\log y_t^2 = \sigma + h_t + \varepsilon_t$ $(1 - L)^d x_t = \eta_t$	$\log y_t^2 = \sigma + h_t + \varepsilon_t$ $(1 - \phi L)(1 - L)^d x_t = \eta_t$
Parámetros estimados	$\hat{\sigma}_\eta^2 = 0.0906$	$\hat{\sigma}_\eta^2 = 0.0155$
	-	$\hat{\phi} = 0.6632$
	$\hat{d} = 0.7538$	$\hat{d} = 0.7035$
	$\hat{\sigma}_\varepsilon = 1.5112$	$\hat{\sigma}_\varepsilon = 1.5485$
$Q_2(10)$	17.89	15.50
$Q_2(100)$	117.67	113.39

Nota: $Q_2(K)$ denota el estadístico de Box-Ljung para contrastar incorrelación hasta de orden K en los cuadros de las observaciones estandarizadas.

Como alternativa al método de estimación pseudo-máximo verosímil de Whittle, se han propuesto recientemente otros métodos de estimación paramétricos. Hsu y Breidt (1997), por ejemplo, proponen un método de estimación bayesiano mediante muestreo de Gibbs que permite obtener la distribución a posteriori de los parámetros del modelo LMSV y la estimación suavizada de la volatilidad. So (1999) estudia el problema de estimación de modelos ARFIMA(0[¥[0) "contaminados" por un ruido gaussiano, y propone técnicas bayesianas de estimación MCMC (Markov Chain Monte Carlo) que podrían aplicarse al modelo LMSV. Aunque las propiedades de estos métodos sean óptimas, estas técnicas son computacionalmente muy costosas y no son fáciles de extender a modelos más complicados. De hecho, cuando el tamaño muestral es grande o el proceso ARFIMA contiene componentes autorregresivos o de medias móviles, existen serios problemas en la aplicación de estos métodos. Wright (1999) propone un estimador del modelo LMSV esta-

GRÁFICO 5
Rendimientos absolutos diarios del índice IBEX-35 junto con las estimaciones suavizadas de la volatilidad (en negrita) para el modelo LMSV básico.



cionario basado en el método de los momentos generalizados (GMM) y demuestra que, en el rango $1]2 \setminus ¥ \setminus 1]4$, dicho estimador es $\sqrt{t}$ -consistente y asintóticamente normal. Sin embargo, la utilidad práctica de este resultado puede ser limitada porque los valores del parámetro ¥ en series reales suelen estar cerca de la frontera de la no estacionariedad ($¥ = 1]2$), donde la distribución asintótica es desconocida. Además, las simulaciones de Wright (1999) muestran, en primer lugar, que la distribución asintótica no es una buena aproximación en muestras finitas, y en segundo lugar, que el estimador GMM se compara desfavorablemente con el estimador de Whittle, ya que ambos tienen varianzas similares pero este último tiene un sesgo mucho menor. Finalmente, Chan y Petris (2000) proponen un método de estimación del modelo LMSV, basado en representar en el espacio de los estados el proceso ARFIMA truncado en primeras diferencias y maximizar la función de verosimilitud resultante mediante el filtro de Kalman. Sin embargo, aún no hay resultados sobre las propiedades de este estimador.

Como alternativa a los métodos de estimación paramétricos que acabamos de describir, Deo y Hurvich (2001) han propuesto un estimador semiparamétrico para el parámetro de memoria, ¥, basado en la idea original de Geweke y Porter-Hudak (1983) para modelos ARFIMA.

Este estimador se obtiene a partir de un modelo de regresión para el logaritmo del periodograma de la serie transformada, $^{1}{}_{\mu} = \text{-}^{-}(^{o}{}^{2}_{\mu})$, utilizando sólo las ® primeras frecuencias en torno al origen. Bajo ciertas condiciones sobre ®, Deo y Hurvich (1998) derivan las expresiones del sesgo y de la varianza asintótica del estimador, y demuestran que, cuando ® verifica ciertas condiciones, el estimador es asintóticamente normal, pero con una tasa de convergencia $^{\pm}\overline{\text{®}}$ inferior a la $^{\pm}\overline{\text{t}}$ típica de los estimadores paramétricos. Aunque estos resultados sobre la estimación del parámetro ¥ son importantes, no se conoce aún ningún resultado sobre la estimación de los otros parámetros del modelo ni sobre la estimación de la volatilidad.

En cuanto a la predicción de modelos LMSV, Harvey (1998) proporciona una expresión para predecir valores futuros de la volatilidad en el modelo LMSV básico, pero ni las propiedades analíticas ni empíricas de dichas predicciones son conocidas.

6.3. Otros modelos con memoria larga en volatilidad

En esta sección se describen brevemente otros modelos con memoria larga para la volatilidad que no han tenido mucha repercusión en aplicaciones empíricas.

En primer lugar, Ding y Granger (1996) proponen un modelo en el que la volatilidad es una suma ponderada de n componentes, cada una de las cuales es la varianza condicionada de un modelo GARCH(1,1) con parámetros $\{^{+}{}_{a}[\ ,\ {}_{a}\}$. Bajo ciertas condiciones sobre dichos parámetros, se demuestra que cuando n tiende a infinito, el modelo propuesto admite una representación ARCH(Q) cuyos coeficientes decrecen hiperbólicamente hacia cero; por tanto, la varianza condicionada tiene memoria larga. Este modelo es similar al modelo FIGARCH de Baillie, Bollerslev y Mikkelsen (1996).

Robinson y Za;aroni (1997) y Za;aroni (1997) proponen una nueva clase de modelos, denominados modelos MA no lineales con memoria larga, en los que la volatilidad, $;_{\mu}$, se modeliza mediante una estructura infinita de retardos de la propia perturbación B_{μ}. Este modelo es capaz de producir memoria larga en la serie de los cuadrados. Za;aroni (1997) propone un método de estimación aproximado de Whittle en el dominio de las frecuencias para este modelo y demuestra que estos estimadores son consistentes y asintóticamente normales.

Posteriormente, Robinson y Za;aroni (1998) proponen un modelo SV no lineal con dos perturbaciones, alternativo al modelo LMSV de Harvey (1998) y Breidt, Crato y de Lima (1998), y demuestran que, bajo ciertas condiciones sobre los momentos de B_μ, la serie de los cuadrados puede tener memoria larga. Para estimar el modelo, los autores proponen un estimador pseudo-máximo verosímil en el dominio de las frecuencias, y argumentan que no sería difícil establecer la consistencia y la distribución asintótica de dicho estimador, pero no presentan ninguna demostración formal de tales resultados.

Finalmente, Beran y Ocker (2001) utilizan el modelo SEMIFAR, originalmente propuesto por Beran y Ocker (1999), para representar la volatilidad de varias series de rendimientos financieros y encuentran una fuerte evidencia de memoria larga en la componente estocástica de la volatilidad para casi todas las series analizadas.

7. Consideraciones finales

En este trabajo se han descrito, en primer lugar, las propiedades de series con memoria larga tanto en la media como en la varianza condicionada. Se ha dedicado especial atención a los procesos ARMA fraccionalmente integrados o procesos ARFIMA, por ser éstos los más utilizados en la modelización de la memoria larga en media y ser a su vez la base sobre la que se construyen gran parte de los modelos con memoria larga en la varianza condicionada. Se han revisado los principales métodos de estimación de modelos con memoria larga en media y varianza y se han discutido las ventajas e inconvenientes de cada uno de ellos. Con respecto a los contrastes de hipótesis, se han presentado los contrastes más utilizados en aplicaciones empíricas para discernir entre procesos con memoria corta I(0) y procesos con memoria larga I(¥), y aquellos que contrastan la hipótesis de raíz unitaria (¥ = 1) frente a la memoria larga.

Aunque en este artículo nos hemos centrado fundamentalmente en la revisión de modelos univariantes con memoria larga, existe una extensa literatura relacionada con modelos multivariantes. De hecho, aunque nosotras los hayamos particularizado para el caso univariante, algunos de los artículos mencionados anteriormente como, por ejemplo, Robinson (1995a), Hosoya (1997) y Lobato y Robinson (1998), presentan resultados para series multivariantes. Recientemente, Lobato (1999) ha propuesto una generalización del estimador bietápico propuesto por Robinson (1995b) al caso multivariante y ha probado su

normalidad asintótica. En el contexto de modelos heteroscedásticos multivariantes con memoria larga, Bollerslev y Jubinski (1999) consideran la modelización conjunta de la volatilidad y el volumen de negocio. Sin embargo, Lobato y Velasco (2000) muestran que el estimador propuesto por Bollerslev y Jubinski (1999) tiene propiedades desconocidas y proponen un estimador alternativo para el que prueban un teorema central del límite. Dicho estimador es una modificación del propuesto por Lobato (1999) que puede aplicarse a series no estacionarias mediante la introducción del procedimiento de *tapering* de Velasco (1999b). Alternativamente, Ray y Tsay (2000) analizan la presencia de memoria larga común entre las volatilidades de diferentes grupos de precios componentes del índice S&P500 mediante un modelo paramétrico LMSV multivariante. La estimación del parámetro de memoria larga se realiza mediante dos métodos alternativos que proporcionan resultados semejantes: el estimador GPH y el estimador pseudo-máximo verosímil espectral de Whittle. Brunetti y Gilbert (2000) estiman un modelo FIGARCH bivariante para la volatilidad de un activo en dos mercados financieros distintos y encuentran un orden de integración fraccional común en ambas volatilidades y confirman que éstas están fraccionalmente cointegradas. Finalmente, Robinson y Yajima (2002) desarrollan métodos semiparamétricos para analizar la existencia de cointegración en sistemas fraccionalmente integrados estacionarios. Esta relación de trabajos sobre series multivariantes con memoria larga no pretende ser detallada sino dar algunas pinceladas de los trabajos más recientes en la materia. Por supuesto, las referencias contenidas en los artículos mencionados ayudarán a tener una panorámica más amplia de esta área.

A pesar de la extensa literatura existente sobre modelos con memoria larga, aún son muchos los temas pendientes de investigación. Por ejemplo, en relación a la modelización de la posible existencia de memoria larga en la varianza condicionada de las series financieras, aún se desconocen las propiedades dinámicas de algunos de estos modelos. Tampoco se conoce, para muchos de ellos, la distribución asintótica de los estimadores propuestos y, por tanto, se desconocen también las propiedades de los contrastes de memoria larga que de ellos se derivan. Podría ser interesante también analizar las propiedades de las correlaciones muestrales de los cuadrados de series heteroscedásticas con memoria larga. Ray (1993a) señala que las autocovarianzas muestrales de los cuadrados tienen sesgo negativo, incluso en tamaños de muestras muy grandes, y que este hecho puede explicar porqué habitualmente

las autocorrelaciones muestrales de dicha serie, a pesar de tener una tasa de decrecimiento muy lenta, suelen tomar valores dentro de las bandas de confianza de incorrelación.

Otros temas aún inexplorados son el problema de la predicción en modelos con memoria larga para la varianza condicionada o la posibilidad de incluir componentes estacionales en dichos modelos con el fin de recoger el comportamiento periódico que presentan, por ejemplo, las series de datos financieros *intra-day*. Éstos y otros temas, dejan abiertas numerosas posibilidades de investigación en las que seguir avanzando.

Referencias

Abramowitz, M. y N.C. Stegun (1970), *Handbook of Mathematical Functions,* Dover Publications Inc., New York.

Agiakloglou, C. y P. Newbold (1994): "Lagrange multiplier tests for fractional di ; erence", *Journal of Time Series Analysis* 15, pp. 253-262.

Agiakloglou, C., P. Newbold y M. Wohar (1993): "Bias in an estimator of the fractional di ; erence parameter", *Journal of Time Series Analysis* 14, pp. 235-246.

Andersen, T.G. y T. Bollerslev (1997): "Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns", *Journal of Finance* LII, pp. 975-1005.

Andersen, T.G., T. Bollerslev, F.X. Diebold y P. Labys (2001): "The distribution of realized exchange rate volatility", *Journal of the American Statistical Association* 96, pp. 42-55.

Baillie, R.T. (1996): "Long memory processes and fractional integration in econometrics", *Journal of Econometrics* 73, pp. 5-59.

Baillie, R.T. y T. Bollerslev (1994): "The long memory property of forward premium", *Journal of International Money and Finance* 13, pp. 565-571.

Baillie, R.T., T. Bollerslev y H.O. Mikkelsen (1996): "Fractionally integrated autoregressive conditional heteroskedaticity", *Journal of Econometrics* 74, pp. 3-30.

Baillie, R.T., C.F. Chung y M.A. Tieslau (1996): "Analysing inflation by the fractionally integrated ARFIMA-GARCH model", *Journal of Applied Econometrics* 11, pp. 23-40.

Beran, J. (1994), *Statistics for Long-memory Processes,* Chapman and Hal, New York.

Beran, J. y D. Ocker (1999): "SEMIFAR forecasts, with applications to foreign exchange rates", *Journal of Statistical Planning and Inference* 80, pp. 137-153.

Beran, J. y D. Ocker (2001): "Volatility of stock-market indexes. An analysis based on SEMIFAR models", *Journal of Business and Economic Statistics* 19, pp. 103-116.

Bloomfield, P. (1976), *Fourier Analysis of Time Series: An Introduction,* Wiley, New York.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics* 31, pp. 307-327.

Bollerslev, T. (1988): "On the correlation structure for the generalized autoregressive conditional heteroskedastic process", *Journal of Time Series Analysis* 9, pp. 121-131.

Bollerslev, T. y D. Jubinski (1999): "Equity trading volume and volatility: latent information arrivals and common long-run dependencies", *Journal of Business and Economic Statistics* 17, pp. 9-21.

Bollerslev, T. y H.O. Mikkelsen (1996): "Modeling and pricing long memory in stock market volatility", *Journal of Econometrics* 73, pp. 151-184.

Bollerslev, T. y H.O. Mikkelsen (1999): "Long-term equity anticipation securities and stock market volatility dynamics", *Journal of Econometrics* 92, pp. 75-99.

Breidt, F.J., N. Crato y P.J.F. de Lima (1998): "The detection and estimation of long-memory in stochastic volatility", *Journal of Econometrics* 83, pp. 325-348.

Brockwell, P.J. y R.A. Davis (1991), *Time Series: Theory and Methods,* 2ª ed., Springer Verlag, Berlin.

Brodsky, J. y C.M. Hurvich (1999): "Multi-step forecasting for long-memory processes", *Journal of Forecasting* 18, pp.59-75.

Brunetti, C. y C.L. Gilbert (2000): "Bivariate FIGARCH and fractional cointegration", *Journal of Empirical Finance* 7, pp. 509-530.

Chan, N.H. y W. Palma (1998): "State space modeling of long memory processes", *Annals of Statistics* 26, pp. 719-740.

Chan, N.H. y G. Petris (2000): "Long memory stochastic volatility: a bayesian approach", *Communications in Statistics. Theory and Methods* 29, pp. 1367-1378.

Chen, G., B. Abraham y S. Peiris (1994): "Lag window estimation of the degree of di ; erencing in fractionally integrated time series models", *Journal of Time Series Analysis* 15, pp. 473-487.

Cheung, Y.W. (1993a): "Long memory in foreign-exchange rates", *Journal of Business and Economic Statistics* 11, pp. 93-101.

Cheung, Y.W. (1993b): "Tests for fractional integration: a Monte Carlo investigation", *Journal of Time Series* 14, pp. 331-345.

Cheung, Y.W. y F.X. Diebold (1994): "On maximum likelihood estimation of the di ; erencing parameter of fractionally integrated noise with unknown mean", *Journal of Econometrics* 62, pp. 301-316.

Chung, C.F. (1994): "A note on calculating the autocovariances of the fractionally integrated ARMA models", *Economics Letters* 45, pp. 293-297.

Chung, C.F. y R.T. Baillie (1993): "Small sample bias in conditional sum of squares estimators of fractionally integrated ARMA models", *Empirical Economics* 18, pp. 791-806.

Crato, N y P.J.F. de Lima (1994): "Long-range dependence in the conditional variance of stock returns", *Economics Letters* 45, pp. 281-285.

Crato, N. y B.K. Ray (1996): "Model selection and forecasting for long-range dependent processes", *Journal of Forecasting* 15, pp. 107-125.

Crato, N. y H.M. Taylor (1996): "Stationary persistent time series misspecified as nonstationary ARIMA", *Statistical Papers* 37, pp. 215-223.

Dacorogna, M.M., U.A. Müller, R.J. Nagler, R.B. Olsen y O.V. Pictet (1993): "A geographical model for the daily and weekly seasonal volatility in the foreign exchange market", *Journal of International Money and Finance* 12, pp. 413-438.

Dahlhaus, R. (1989): "E>cient parameter estimation for self-similar processes", *The Annals of Statistics* 17, pp. 1749-1766.

Davies, R.B. y D.S. Harte (1987): "Tests for Hurst e;ect", *Biometrika* 74, pp. 95-101.

Delgado, M. y P.M. Robinson (1993): "New methods for the analysis of long-memory time series: application to Spanish inflation", *Journal of Empirical Finance* 1, pp. 83-106.

Deo, R. y C.M. Hurvich (2001): "On the log-periodogram regression estimator of the memory parameter in long memory stochastic volatility models", *Econometric Theory* 17, pp. 686-710.

Diebold, F.X. y P. Lindner (1996): "Fractional integration and interval prediction", *Economics Letters* 50, pp. 305-313.

Diebold, F.X y G.D. Rudebusch (1989): "Long-memory and persistence in aggregate output", *Journal of Monetary Economics* 24, pp. 189-209.

Diebold, F.X. y G.D. Rudebusch (1991): "On the power of Dickey-Fuller tests against fractional alternatives", *Economics Letters* 35, pp. 155-160.

Ding, Z. y C.W.J. Granger (1996): "Modeling volatility persistence of speculative returns: a new approach", *Journal of Econometrics* 73, pp. 185-215.

Ding, Z., C.W.J. Granger y R.F. Engle (1993): "A long memory property of stock market returns and a new model", *Journal of Empirical Finance* 1, pp. 83-106.

Dolado, J.J. y F. Marmol (1997): "On the properties of the Dickey-Pantula test against fractional alternatives", *Economics Letters* 57, pp. 11-16.

Dolado, J.J., J. Gonzalo y L. Mayoral (1999): "A fractional Dickey-Fuller test", Working Paper 99-29(8), Universidad Carlos III de Madrid.

Eisinga, R., P.H. Franses y M. Ooms (1999): "Forecasting long memory left-right political orientations", *International Journal of Forecasting* 15, pp. 185-199.

Engle, R.F. (1982): "Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation", *Econometrica* 50, pp. 987-1007.

Fox, R. y M.S. Taqqu (1986): "Large sample properties of parameter estimates for strongly dependent stationary gaussian time series", *Annals of Statistics* 14, pp. 221-238.

Franses, P.H. y M. Ooms (1997): "A periodic long memory model for quarterly UK inflation", *International Journal of Forecasting* 13, pp. 117-126.

Geweke, J. y S. Porter-Hudak (1983): "Estimation and application of long memory time series models", *Journal of Time Series Analysis* 4, pp. 221-238.

Gil-Alana, L.A. y P.M. Robinson (1997): "Testing of unit root and other nonstationary hypotheses in macroeconomics", *Journal of Econometrics* 80, pp. 241-268.

Giriatis, L. y D. Surgalis (1990): "A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle´s estimate", *Probability Theory and Related Fields* 86, pp. 87-104.

Gradshteyn, I.S. e I.M. Ryzhik (1980), *Table of Integrals, Series and Products,* Academic Press, San Diego.

Granger, C.W.J. (1966): "The typical spectral shape of an economic variable", *Econometrica* 34, pp. 150-161.

Granger, C.W.J. (1980): "Long memory relationships and the aggregation of dynamic models", *Journal of Econometrics* 14, pp. 227-238.

Granger, C.W.J. (1990): "Aggregation of time series variables: a survey", en Baker, T. y M. H. Pesaran (Editores) *Disaggregation in Econometric Modelling,* Routledge, London y New York, pp. 17-34.

Granger, C.W.J. y R. Joyeux (1980): "An introduction to long-memory time series models and fractional di;erencing", *Journal of Time Series Analysis* 1, pp. 15-29.

Gray, H.L., N-F. Zhang y W.A. Woodward (1989): "On generalized fractional processes", *Journal of Time Series Analysis* 10, pp. 233-257.

Greene, M. y B. Fielitz (1977): "Long-term dependence in common stock returns", *Journal of Financial Economics* 4, pp. 339-349.

Harvey, A.C. (1998): "Long memory in stochastic volatility,", en Knight, J. y S. Satchell (Editores) *Forecasting Volatility in Financial Markets,* Butterworth-Haineman, Oxford, pp. 307-320.

Hassler, U. (1993): "Regression of spectral estimators with fractionally integrated time series", *Journal of Time Series Analysis* 14, pp. 369-380.

Hassler, U. y J. Wolters (1994): "On the power of unit root tests against fractional alternatives", *Economics Letters* 45, pp. 1-5.

Hassler, U. y J. Wolters (1995): "Long memory in inflation rates: international evidence", *Journal of Business and Economic Statistics* 13, pp. 37-45.

Haubrich, J.G. y A.W. Lo (1989): "The sources and nature of long-term memory in the business cycle", NBER Working Paper 2951.

Hosking, J.R.M. (1981): "Fractional di;erencing", *Biometrika* 68, pp. 165-176.

Hosking, J.R.M. (1996): "Asymptotic distributions of the sample mean, autocovariances and autocorrelations of long-memory time series", *Journal of Econometrics* 73, pp. 261-284.

Hosoya, Y. (1996): "The quasi-likelihood approach to statistical inference on multiple time-series with long-range dependence", *Journal of Econometrics* 73, pp. 217-236.

Hosoya, Y. (1997): "A limit theorem for long-range dependence and statistical inference on related models", *The Annals of Statistics* 25, pp. 105-137.

Hsu, N.J. y F.J. Breidt (1997): "Bayesian analysis of long memory stochastic volatility models", Technical report 97-4, Department of Statistics, Iowa State University.

Hurst, H.E. (1951): "Methods for storage capacity of reservoirs", *Trans. Am. Soc. Civil Engineers* 116, pp. 770-799.

Hurvich, C.M. y K.I. Beltrao (1994): "Automatic semiparametric estimation of the memory parameter of a long-memory time series", *Journal of Time Series Analysis* 15, pp. 285-302.

Hurvich, C.M. y J. Brodsky (2001): "Broadband semiparametric estimation of the memory parameter of a long-memory time series using fractional exponential models", *Journal of Time Series Analysis* 22, pp. 221-249.

Hurvich, C.M., R. Deo y J. Brodsky (1998): "The mean square error of Geweke and Porter-Hudak estimator of the memory parameter in a long-memory time series", *Journal of Time Series Analysis* 19, pp. 19-46.

Johnson, N.L. y S. Kotz (1970), *Continuous Uivariate Distributions 2,* Wiley, New York.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt e Y. Shin (1992): "Testing the null hypothesis of stationarity against the alternative of a unit root", *Journal of Econometrics* 54, pp. 159-178.

Lee, D. y P. Schmidt (1996): "On the power of the KPSS test of stationarity against fractionally integrated alternatives", *Journal of Econometrics* 73, pp. 285-302.

Lee, H.S. y C. Amsler (1997): "Consistency of the KPSS unit root test against fractionally integrated alternative", *Economics Letters* 55, pp. 151-160.

Li, W.K. y A.I. McLeod (1986): "Fractional time series modelling", *Biometrika* 73, pp. 217-221.

Lo, A.W. (1991): "Long-term memory in stock market prices", *Econometrica* 59, pp. 1279-1313.

Lobato, I. (1999): "A semiparametric two-step estimator in a multivariate long memory model", *Journal of Econometrics* 90, pp. 129-153.

Lobato, I. y P.M. Robinson (1996): "Averaged periodogram estimation of long memory", *Journal of Econometrics* 73, pp. 303-324.

Lobato, I. y P.M. Robinson (1998): "A nonparametric test for I(0)", *Review of Economic Studies* 65, pp. 475-495.

Lobato, I. y N.E. Savin (1998): "Real and spurious long-memory properties of stock-market data", *Journal of Business and Economic Statistics* 16, pp. 261-283.

Lobato, I. y C. Velasco (2000): "Long memory in stock market trading volume", *Journal of Business and Economic Statistics* 18, pp. 410-427.

Marmol, F. (1998): "Searching for fractional evidence using combined unit root tests", *Working Paper 98-39(19)* Universidad Carlos III de Madrid.

Moulines, E. y P. Soulier (1999): "Broad band log-periodogram regression of time series with long range dependence", *Annals of Statistics* 27, pp. 1415-1439.

Moulines, E. y P. Soulier (2000): "Data-driven order selection for long range dependent time series", *Journal of Time Series Analysis* 21, pp. 193-218.

Nelson, C.R. y C.I. Plosser (1982): "Trends and random walks in macroeconomic time series", *Journal of Monetary Economics* 10, pp. 139-162.

Nelson, D.B. (1991): "Conditional heteroscedasticity in asset returns: a new approach", *Econometrica* 59, pp. 347-370.

Nelson, D.B. y C.Q. Cao (1992): "Inequality constraints in the univariate GARCH model", *Journal of Business and Economic Statistics* 10, pp. 229-235.

Newbold, P. y C. Agiakloglou (1993): "Bias in the sample autocorrelations of fractional noise", *Biometrika* 80, pp. 698-702.

Olmeda, I. (1998): "Algunos resultados sobre memoria de largo plazo en series bursátiles", *Moneda y Crédito* 207, pp. 145-203.

Parke, W.R. (1999): "What is fractional integration?", *The Review of Economics and Statistics* 81, pp. 632-638.

Peiris, M.S. y B.J. Perera (1988): "On prediction with fractionally di;erenced ARIMA models", *Journal of Time Series Analysis* 9, pp. 215-220.

Pérez, A. (2000): "Estimación e identificación de modelos de volatilidad estocástica con memoria larga", Tesis doctoral, Departamento de Economía Aplicada (Estadística y Econometría), Universidad de Valladolid.

Pérez, A. y E. Ruiz (2001): "Finite sample properties of a QML estimator of stochastic volatility models with long memory", *Economics Letters* 70, 157-164.

Porter-Hudak, S. (1990): "An application of the seasonal fractionally differenced model to the monetary aggregates", *Journal of the American Statistical Association* 85, pp. 338-344.

Priestley, M.B. (1981), *Spectral Analysis and Time Series,* Academic Press, London.

Ray, B.K. (1993a): "Long-range forecasting of IBM product revenues using a seasonal fractionally di;erenced ARMA model", *International Journal of Forecasting* 9, pp. 255-269.

Ray, B.K. (1993b): "Modeling long-Memory processes for optimal long-range prediction", *Journal of Time Series Analysis* 14, pp. 511-525.

Ray, B.K. y R.S. Tsay (2000): "Long-range dependence in daily stock volatilities", *Journal of Business and Economic Statistics* 18, pp. 254-262.

Reisen, V.A. (1994): "Estimation of the fractional di;erence parameter in the ARIMA(p,d,q) model using the smoothed periodogram", *Journal of Time Series Analysis* 15, pp. 335-350.

Reisen, V.A. y S. Lopes (1999): "Some simulations and applications of forecasting long-memory time series models", *Journal of Statistical Planning and Inference* 80, pp. 269-287.

Robinson, P.M. (1978): "Statistical inference for random coe>cient autoregressive model", *Scandinavian Journal of Statistics* 5, pp. 163-168.

Robinson, P.M. (1991): "Testing for strong serial correlation and dynamic conditional heteroskedaticity in multiple regression", *Journal of Econometrics* 47, pp. 67-84.

Robinson, P.M. (1994a): "E>cient tests of nonstationary hypotheses", *Journal of the American Statistical Association* 89, pp. 1420-1437.

Robinson, P.M. (1994b): "Time series with strong dependence", en Sims, C. A. (ed.) *Advances in Econometrics: Sixth World Congress* (Volumen 1), Cambridge University Press, Cambridge, pp. 47-96.

Robinson, P.M. (1995a): "Log-periodogram regression of time series with long-run dependence", *Annals of Statistics* 23, pp. 1048-1072.

Robinson, P.M. (1995b): "Gaussian semiparametric estimation of long-range dependence", *Annals of Statistics* 23, pp. 1630-1661.

Robinson, P.M. y Y. Yajima (2002): "Determination of cointegrating rank in fractional systems", *Journal of Econometrics* 106, pp. 217-241.

Robinson P.M. y P. Za;aroni (1997): "Modelling nonlinearity and long memory in time series", en Cutler, C.D. y D.T. Kaplan (eds.) *Nonlinear Dynamics and Time Series, Fields Institute Communications* 11, pp. 161-170.

Robinson P.M. y P. Za;aroni (1998): "Nonlinear time series with long memory: a model for stochastic volatility", *Journal of Statistical Planning and Inference* 68, pp. 359-371.

Smith, A.A., F. Sowell y S.E. Zin (1997): "Fractional integration with drift: estimation in small samples", *Empirical Economics* 22, pp. 103-116.

Smith, J., N. Taylor y S. Yadav (1997): "Comparing the bias and misspecification in ARFIMA models", *Journal of Time Series Analysis* 18, pp. 507-527.

So, M.K.P. (1999): "Time series with additive noise", *Biometrika* 86, pp. 474-482.

Sowell, F. (1990): "The fractional unit root distribution", *Econometrica* 58, pp. 495-505.

Sowell, F. (1992a): "Maximum likelihood estimation of stationary univariate fractionally integrated time series models", *Journal of Econometrics* 53, pp. 165-188.

Sowell, F. (1992b): "Modelling long-run behaviour with the fractional ARIMA model", *Journal of Monetary Economics* 29, pp. 277-302.

Sutcli;e, A. (1994): "Time-series forecasting using fractional di;erencing", *Journal of Forecasting* 13, pp. 383-393.

Taylor, S. (1986), *Modelling Financial Time Series,* Wiley, New York.

Teverovsky, V., M. S. Taqqu y W. Willinger (1999): "A critical look at Lo's modified R/S statistic", *Journal of Statistical Planning and Inference* 80, pp. 211-227.

Tiao, G.C. y R.S. Tsay (1994): "Some advances in non-linear and adaptive modeling time series", *Journal of Forecasting* 13, pp. 109-131.

Velasco, C. (1999a): "Gaussian semiparametric estimation of nonstationary time series", *Journal of Time Series Analysis* 20, pp. 87-127.

Velasco, C. (1999b): "Non-stationary log-periodogram regression", *Journal of Econometrics* 91, pp. 325-371.

Velasco, C. (2000): "Non-gaussian log-periodogram regression", *Econometric Theory* 16, pp. 44-79.

Velasco, C y P.M. Robinson (2000): "Whittle pseudo-maximum likelihood estimates of nonstationary time series", *Journal of the American Statistical Association* 95, pp. 1229-1243.

Whittle, P. (1962): "Gaussian estimation in stationary time series", *Bull. Internat. Statistical Institute* 39, pp. 105-129.

Wright, J.H. (1999): "A new estimator of the fractionally integrated stochastic volatility model", *Economics Letters* 63, pp. 295-303.

Yajima, Y. (1985): "On estimation of long-memory time series models", *Australian Journal of Statistics* 27, pp. 303-320.

Za ; aroni, P. (1997): "Gaussian estimation of long-range dependent volatility in asset prices", STICERD Econometrics Discussion Paper EM/97/329, London School of Economics.

Abstract

This paper provides a review of time series models with long memory in the mean and conditional variance, with special attention to Fractionally Integrated ARMA processes (ARFIMA) and fractionally integrated GARCH and SV processes. Their more important properties are reviewed and its application to model economic and financial time series is discussed. The main estimation methods and tests proposed in the literature for the long memory property are also reviewed. Finally, this paper reviews the main results on prediction of future values of long memory time series.

Keywords: Fractional integration, persistency, frequency domain, spectral density, volatility.

Recepción del original, junio de 2001
Versión final, enero de 2002