

MODELOS AUTORREGRESIVOS PARA LA VARIANZA CONDICIONADA HETEROSCEDASTICA (ARCH)

Marc Saez Zafra¹
Jorge V. Pérez Rodríguez²

1.- Departamento de Economía. Universidad Pompeu Fabra.

2.- Departamento de Econometría. Universidad de Barcelona.

RESUMEN

En el presente trabajo se introduce al lector en los modelos autorregresivos para la varianza condicionada heterocedástica, incidiendo en los problemas que plantean los esquemas más sencillos y sugiriendo diversas soluciones. Se describe el concepto, las hipótesis y los modelos que explican la varianza condicionada en el tiempo desde diversas perspectivas del análisis estadístico: relación lineal o no lineal entre las variables y métodos de estimación de los parámetros. Finalmente, se discuten diversos contrastes que permiten escoger entre diversas especificaciones alternativas.

ABSTRACT

In this paper the Autoregressive Conditional Heteroskedasticity (ARCH) models are introduced to the reader. The problems implied by their different parameterizations are pointed out and the corresponding solutions are suggested. Several statistical perspectives are used to explain the concept, the assumptions and the models that explain the temporal behaviour of the conditional variances. The paper exposed lineal and non-lineal relationships and several methods of estimation of the parameters of the models. Finally, some tests that allow to choose between alternative specifications are described.

Palabras Clave: Modelos ARCH, varianzas condicional e incondicional, variabilidad temporal.

Keywords: ARCH models, conditional and unconditional variances, time-varying factors.

1.- INTRODUCCIÓN.

En la Teoría Económica es práctica habitual suponer que los agentes económicos no solamente responden y actúan ante los cambios en los valores medios de las variables económicas, sino que también tienen en cuenta otros momentos de orden superior para tomar sus decisiones. Así, la teoría financiera considera no sólo a la rentabilidad sino también, y quizás primordialmente, a la volatilidad (varianza) como determinantes de las decisiones de cartera de los agentes.

Contrariamente, en la econometría estos momentos siempre han sido discriminados. En la econometría tradicional, por ejemplo, se ha supuesto que la varianza de la predicción es constante en el tiempo, ignorando que la incertidumbre y la aleatoriedad inherentes en los fenómenos económicos la modifican continuamente. Del mismo modo sucede en el análisis de series temporales, en el que siempre se ha exigido la estacionariedad de la varianza.

Sin embargo los modelos estadísticos modernos, por ejemplo los que realizan la valoración de los precios de los activos financieros, utilizan como elemento diferenciador los momentos condicionados y no condicionados de orden superior al primero, o *r*-ésimos. Las técnicas que se utilizan para modelizar los momentos de segundo orden, o más concretamente la heteroscedasticidad, resultan una ampliación del análisis de series temporales, popularizadas para los momentos de primer orden por **Box y Jenkins (1976)**. Estos modelos, en los que la varianza de la predicción puede cambiar en el tiempo, se denominan modelos Autorregresivos Heteroscedásticos Condicionados (ARCH) y fueron propuestos originariamente por **Engle (1982)**.

La amplia difusión de los modelos ARCH en trabajos empíricos, en especial en economía financiera, ha radicado en su éxito a la hora de sintetizar los rasgos que caracterizan el comportamiento dinámico a corto plazo de las variables involucradas, a saber agrupación ('clustering') de las volatilidades y la existencia de distribuciones incondicionales de probabilidad altamente leptocúrticas. Al tiempo que los modelos ARCH se han ido popularizando en la última década, también lo han hecho las panorámicas sobre los mismos, véase **Nijman y Palm (1991)**, **Bera y Higgins (1992)**, y sobretodo **Bollerslev, Chou y Kroner (1992)**, por sólo citar algunos ejemplos. En castellano, el lector interesado puede recurrir al excelente artículo de **Novalés y Gracia-Díez (1993)**. En la misma línea, el presente trabajo pretende introducir al lector en estos modelos, incidiendo en los problemas que plantean los esquemas más sencillos y sugiriendo diversas soluciones. Así, se va a tratar de describir el concepto, las hipótesis y los modelos que explican la varianza condicionada en el tiempo desde diversas perspectivas del análisis estadístico: relación lineal o no lineal entre las variables y métodos de estimación de los parámetros.

2.- MODELOS AUTORREGRESIVOS CONDicionados HETEROCEDASTICOS.

Los modelos ARCH son una clase de procesos estocásticos que están caracterizados porque, aún teniendo media nula y estando serialmente incorrelacionados, poseen una varianza condicionada que no es constante en el tiempo.

En el modelo ARCH más sencillo (Engle, 1982) se parte de un esquema autorregresivo de orden uno (aunque para su formulación pueden utilizarse otros modelos ARMA más complejos, véase Weiss (1984)):

$$y_t = \phi_1 y_{t-1} + \varepsilon_t$$

en el que ε_t es un ruido blanco ($\varepsilon_t \approx \text{i.i.d. } N(0, \sigma^2)$).

Si se representa por $E_t[\cdot]$ el operador matemático esperanza condicionada a la información existente, la media condicionada del proceso estocástico y_t es,

$$E_t[y_t] = \phi_1 y_{t-1}$$

mientras que la no condicionada (la media del AR(1) en este caso) es cero ya que no existe término independiente,

$$E[y_t] = 0$$

Por otro lado, la varianza no condicionada de y_t (del AR(1)) es igual a

$$\text{Var}[y_t] = \frac{\sigma^2}{1 - \phi_1^2}$$

mientras que la varianza condicionada de y_t es,

$$\text{Var}_t[y_t] = E_t[y_t - E_t[y_t]]^2 = E_t[\phi_1 y_{t-1} + \varepsilon_t - \phi_1 y_{t-1}]^2 = E[\varepsilon_t^2] = \sigma_\varepsilon^2$$

Así, la distribución de y_t condicionada a la información existente es

$$y_t / \Psi_{t-1} \sim N(0, h_t)$$

donde, Ψ_{t-1} es el conjunto de información disponible en $t-1$, y con h_t^2 se representa la varianza de y_t , dependiente del tiempo en nuestro caso.

Así pues, el modelo AR(1)-ARCH(1) la varianza condicionada de y_t sigue un polinomio autorregresivo de orden uno:

$$\begin{aligned} y_t &= \phi_1 y_{t-1} + \varepsilon_t \\ E_{t-1}[\varepsilon_t] &= 0 \\ h_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \end{aligned} \quad [1]$$

siendo α_0 y α_1 parámetros desconocidos.

Los errores ARCH también pueden obtenerse a partir de un modelo de regresión con perturbaciones heteroscedásticas,

$$y_t = x_t \beta + u_t$$

donde, y_t es la variable dependiente, x_t es un conjunto de variables exógenas y endógenas retardadas incluidas en el conjunto de información ψ_{t-1} , y, β es un vector de parámetros desconocidos.

La distribución de la variable dependiente condicionada a toda la información existente es

$$y_t / \psi_{t-1} \sim N(x_t \beta, h_t^2)$$

en el que se ha supuesto la normalidad de la variable dependiente

Definiendo ε_t , error en t , como la diferencia entre el valor observado y el valor estimado de la variable dependiente en t ,

$$\varepsilon_t = y_t - x_t \beta$$

la varianza heteroscedástica, h_t^2 , puede ahora formularse como

$$h_t^2 = h(\varepsilon_{t-1}, \dots, \varepsilon_{t-p}, x_t, x_{t-1}, \dots, x_{t-p}, \alpha)$$

que supone incluir todo el conjunto de información en $t-1$, o bien de una forma más sencilla como

$$\varepsilon_t^2 = h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 = \alpha_0 + \alpha(L) \varepsilon_{t-1}^2 \quad [2]$$

donde, α_0 es el término independiente del polinomio autorregresivo, relacionado con la varianza no condicionada, y, $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_p L^p$. De esta forma, un ARCH(1) no es más que

$$\begin{aligned}y_t &= x_t \beta + \varepsilon_t \\ E[\varepsilon_t] &= 0 \\ E_{t-1}[\varepsilon_t^2] &= h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2\end{aligned}$$

donde, $E_{t-1}[\varepsilon_t^2]$ es la varianza condicionada a la información en $t-1$.

Hipótesis del modelo ARCH

Del artículo de Engle (1982) se pueden extraer tres hipótesis iniciales que caracterizan al modelo ARCH(p).

Dependencia lineal de los valores de ε_t

En esta hipótesis se hace referencia conjuntamente tanto a la dependencia lineal de ε_t^2 respecto a sus errores pasados y a la forma funcional lineal para la varianza condicionada.

Normalidad condicionada del error

Se supone que las observaciones generadas a partir de un modelo ARCH se distribuyen condicionalmente como una normal. Esta hipótesis permite aplicar la siguiente propiedad. Para una variable aleatoria X que está normalmente condicionada a la información existente, con media nula y varianza constante, el segundo r -ésimo momento de la misma se puede expresar como,

$$E[X^{2r}] = \sigma^{2r} \prod_{j=1}^r (2j-1)$$

Así, cuando la distribución condicionada de ε_t es normal y la varianza heterocedástica sigue un esquema ARCH(1),

$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

con parámetros $\alpha_0 > 0$ y $\alpha_1 > 0$, el segundo m -ésimo momento puede escribirse como

$$E[h_t^{2m} | \Psi_{t-1}] = h_t^{2m} \prod_{j=1}^m (2j-1) = (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2) \prod_{j=1}^m (2j-1)$$

Acotación de las varianzas

Una hipótesis adicional que podemos considerar en la especificación del modelo ARCH es que la varianza ha de ser estrictamente positiva y finita.

Varianza estrictamente positiva

Para asegurar que la varianza condicionada sea estrictamente positiva en todos los valores o realizaciones pasadas del error, ε_i , el modelo ARCH necesita que el espacio de parámetros esté acotado a ser estrictamente positivo. Esto es, todos los valores de α_i en [2] deben ser positivos, $\alpha_i > 0$, $i=1, \dots, p$.

Así, en el modelo ARCH(1) $\alpha_0 > 0$ y $\alpha_1 > 0$. Por otra parte, α_0 coincide con el valor de la varianza no condicionada que es constante.

Varianza finita o estacionaria

Para que las varianzas sean estacionarias y finitas, es decir, para que su función de densidad conjunta sea constante en el tiempo, la acotación de la varianza requiere que en [2] se cumpla el **teorema 1** de Engle (1982) el cual determina que el 2r-ésimo momento de un ARCH(1) debe verificar que

$$\alpha_1^2 \prod_{j=1}^r (2j-1) < 1$$

Así, un modelo ARCH(1) con parámetros estrictamente positivos $\alpha_0 > 0$ y $\alpha_1 > 0$, tendrá un segundo momento condicionado positivo. Pero para que este momento esté acotado, es decir para que la varianza sea estacionaria y finita, se requiere además que $\alpha_1 < 1$.

Partiendo de [1],

$$h_t^2 = \varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 = E[h_t^2] = \alpha_0 + \alpha_1 E[h_{t-1}^2]$$

$$E[h_t^2] = \frac{\alpha_0}{1 - \alpha_1}$$

Un proceso ARCH lineal de orden p será estacionario en covarianza si y sólo si, las raíces de la ecuación característica de h_t^2 , caen fuera del círculo unidad (Engle, 1982). Esta condición implica

$$\sum_{i=1}^p \alpha_i < 1$$

que no es más que su varianza incondicional, si el ARCH(1) es estacionario ($\alpha^1 < 1$). Mientras que la varianza incondicional de un ARCH(p) estacionario es,

$$Var[\varepsilon_t] = \sigma_\varepsilon^2 = \frac{\alpha_0}{1 - \sum_{i=1}^p \alpha_i}$$

Obsérvese que,

$$h_t^2 - \sigma_\varepsilon^2 = \alpha_1 (\varepsilon_{t-1}^2 - \sigma_\varepsilon^2)$$

es decir, la varianza condicional de h_t^2 es mayor que su varianza incondicional, σ_ε^2 , cuando se producen 'sorpresas' altas (positivas o negativas).

Otros momentos incondicionados de orden superior pueden ser derivados a partir de las tres hipótesis vistas previamente (Engle, 1982). Bajo la hipótesis de normalidad condicionada los momentos de ε_t de orden impar son nulos por simetría. Mientras que los de orden par pueden ser derivados definiendo un vector w_t de variables retardadas, $w_t' = (\varepsilon_{t-1}^2, \dots, \varepsilon_{t-p}^2)$, y resolviendo la siguiente ecuación:

$$E[w_t / \psi_{t-1}] = b + A w_{t-1}$$

en la que,

$$b' = [\alpha_0, 0, \dots, 0]'$$

$$A = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_p & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

Aplicando sustituciones sucesivas

$$E[w_t / \Psi_{t-k}] = (I + A + A^2 + \dots + A^{k-1})b + A^k w_{t-k}$$

y obteniendo el valor en el límite se consigue la expresión para los momentos estacionarios no condicionados de ε_t ,

$$\lim_{k \rightarrow \infty} E[w_t / \Psi_{t-k}] = E[w_t] = (I - A)^{-1}b$$

En el ejemplo más simple, si deseamos obtener el cuarto y el segundo momento estacionario no condicionado de un ARCH(1), tal que, $w_t = [\varepsilon_{t-1}^4, \varepsilon_{t-1}^2]'$, aplicando el segundo supuesto, tenemos que

$$E[\varepsilon_{t-1}^4] = (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^2 \prod_{j=1}^2 (2j-1) = 3\alpha_0^2 + 3\alpha_1^2 \varepsilon_{t-1}^4 + 6\alpha_0 \alpha_1 \varepsilon_{t-1}^2$$

$$E[\varepsilon_{t-1}^2] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

entonces

$$E[w_t / \Psi_{t-1}] = \begin{bmatrix} 3\alpha_0^2 \\ \alpha_0 \end{bmatrix} + \begin{bmatrix} 3\alpha_1^2 & 6\alpha_0 \alpha_1 \\ 0 & \alpha_1 \end{bmatrix} w_{t-1}$$

siendo,

$$b' = [3\alpha_0^2 \quad \alpha_0]$$

$$A = \begin{bmatrix} 3\alpha_1^2 & 6\alpha_0 \alpha_1 \\ 0 & \alpha_1 \end{bmatrix}$$

por tanto, en el límite el valor de los momentos estacionarios no condicionados es un vector 2x1 de elementos, tales que,

$$E[w_t] = \begin{bmatrix} \left[\frac{3\alpha_0^2}{(1-\alpha_1)^2} \right] \left[\frac{1-\alpha_1^2}{1-3\alpha_1^2} \right] \\ \left[\frac{\alpha_0}{1-\alpha_1} \right] \end{bmatrix}$$

en el que el elemento inferior es la varianza no condicionada del ARCH(1) y el elemento superior es el producto que obtiene el cuarto momento, si α_1 es distinto de cero (tercera hipótesis).

Es decir el momento de cuarto orden no es más que,

$$3(\text{Var}[\epsilon_t])^2 \frac{1-\alpha_1^2}{1-3\alpha_1^2}$$

momento que es mayor que el de una variable normal dependiente e idénticamente distribuida. Es decir, las colas de la distribución marginal de ϵ_t son más 'gruesas' que la distribución normal. De hecho, despejando la ecuación anterior, los momentos de cuarto orden sólo existirán si $\alpha < .557$

$$E[\epsilon_t^4] < \infty$$

$$\alpha_1 < 3^{-\frac{1}{2}} = 0.557$$

y los momentos de octavo orden

$$E[\epsilon_t^8] < \infty$$

$$\alpha_1 < 105^{-\frac{1}{4}} = 0.3123$$

3.- OTROS MODELOS ARCH.

En la extensa literatura sobre los modelos ARCH se encuentran otros tipos de especificaciones, cada una de las cuales hace referencia a una de las hipótesis planteados en el modelo: dependencia, normalidad condicionada y forma funcional.

Hipótesis de dependencia de los valores pasados de ϵ^2_t

El modelo ARCH propuesto por Engle (1982) supone que la varianza condicionada, h^2_t , sólo depende de los valores pasados de ϵ^2_t . Sin embargo, h^2_t puede no depender exclusivamente de los errores pasados del modelo, sino de cualquiera de los elementos o variables incluidas en el conjunto de información Ψ_{t-1} .

Desde este punto de vista, **Weiss (1984 y 1986)** sugiere algunos tipos de procesos ARCH donde h_t^2 es una función de los valores retardados de y_t , de las predicciones de y_t y de otras variables exógenas que se encuentran dentro del conjunto de información. **Bollerslev (1986)** señala la posibilidad de que la varianza condicionada h_t^2 sea, además, una función de los valores retardados de ella misma ($h_{t-1}^2, \dots, h_{t-p}^2$), dando lugar a un tipo de modelos que se conocen como ARCH Generalizados o GARCH de orden (p,q). Desde una perspectiva multivariante **Engle, Granger y Kraft (1984)** y **Granger, Robins y Engle (1984)** indican que la varianza condicionada puede estar relacionada no sólo con los valores retardados de otras variables sino también con sus varianzas.

3.1.- MODELOS ARCH LINEALES

Los modelos lineales alternativos para los procesos de varianza condicionada heteroscedástica pueden ser sintetizados, según la literatura al uso, en cinco grandes especificaciones: modelo GARCH(p,q), modelo ARCH-M, modelo ARCH vectorial, modelo ARCH de factores latentes y modelo estructural de series temporales con errores ARCH.

Modelo GARCH(p,q)

El modelo GARCH (p,q) fue propuesto por **Bollerslev (1986)**. La ecuación que describe este modelo es una suma de polinomios: uno autorregresivo de orden p y otro media móvil de orden q, para la varianza heteroscedástica obtenida del ajuste en [1] o en [2]. Así,

$$h_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}^2 = \omega + \alpha(L) \varepsilon_t^2 + \beta(L) h_t^2 \quad [3]$$

en el que, los valores de los parámetros del modelo deben ser positivos (es decir, $\omega > 0$, $\alpha > 0$ y $\beta > 0$) para evitar que las varianzas sean negativas, las raíces del polinomio $\beta(L)$ deben estar fuera del círculo unitario (como en un modelo ARMA (p,q)), y la varianza no condicionada (ω) debe ser constante para asegurar la estacionariedad en covarianza.

El modelo GARCH(p,q) en [3] puede reescribirse como un modelo ARMA(m,p), tal que

$$\varepsilon_t^2 = \omega + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + \dots + (\alpha_m + \beta_m) \varepsilon_{t-m}^2 - \beta_1 v_{t-1} - \dots - \beta_p v_{t-p} + v_t \quad [4]$$

donde $m = \max(p, q)$, $\alpha_i = 0$ para $i > q$, con ϵ_t y v_t incorrelacionados.

En un modelo GARCH(1,1) tal como,

$$h_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}^2$$

es suficiente que $\alpha_1 + \beta_1 < 1$, para que la varianza heteroscedástica sea estacionaria. En general, en un GARCH(p,q)

$$\sigma_\epsilon^2 = \frac{\alpha_0}{1 - \alpha_1 - \dots - \alpha_q - \beta_1 - \dots - \beta_p}$$

y la condición de estacionariedad es que $\alpha_1 + \dots + \alpha_q + \beta_1 + \dots + \beta_p < 1$.

La condición necesaria y suficiente para la existencia del 2 r-ésimo momento estacionario y finito en el GARCH(1,1) es que,

$$\mu(\alpha, \beta, r) = \sum_{j=0}^r \binom{m}{j} a_j \alpha_1^j \beta_1^{m-j} < 1$$

que resulta de aplicar la fórmula de los momentos con respecto al origen, y en la que por ser una distribución normalmente condicionada se requiere adicionalmente que

$$a_0 = 1$$

$$a_j = \prod_{i=1}^j (2i - 1)$$

Así, el segundo r-ésimo momento no condicionado para ϵ_t puede ser expresado de forma recursiva como

$$E[\epsilon_t^{2r}] = a_r \left[\sum_{n=0}^{r-1} a_n^{-1} E[\epsilon_t^{2n}] \alpha_0^{r-n} \binom{r}{r-n} \mu(\alpha_1, \beta_1, n) \right] \times [1 - \mu(\alpha_1, \beta_1, r)]^{-1}$$

y cuya formulación particular resulta más complicada que la del ARCH.

La mayoría de los modelos empíricos ARCH estimados adoptan el esquema GARCH, en concreto el de un GARCH(1,1). Una explicación a este hecho la proporciona Nelson (1990a). Éste señala que el modelo GARCH(1,1) converge a un modelo de difusión de tiempo continuo cuando la frecuencia de

observación de los datos es 'arbitrariamente' alta (diaria, horaria, etc). Este hecho es muy usual en la teoría financiera, sustentada en ecuaciones diferenciales estocásticas en tiempo continuo, mientras que los datos se observan en tiempo discreto. Una explicación alternativa la proporcionan **Brock, Hsieh y LeBaron (1991)**. Si ϵ_t^2 es lineal, el esquema GARCH(p,q) no es más que la representación parsimoniosa de un proceso infinito de Wold para ϵ_t^2 .

Modelos Integrados en varianza

Si en el esquema más sencillo de generación de datos ARCH, un ARCH (p), se cumple $\alpha_1 + \alpha_2 + \dots + \alpha_p \geq 1$, o en un GARCH (p, q) $\alpha_1 + \alpha_2 + \dots + \alpha_p + \beta_1 + \beta_2 + \dots + \beta_q \geq 1$, entonces existe, al menos, una raíz unitaria en el polinomio definido por su respectiva ecuación característica. Esta situación, que se corresponde con la persistencia en la varianza, es decir una respuesta 'excesivamente' lenta de la misma ante un shock, se denomina integración en varianza y se denota por IARCH(p) o IGARCH(p,q), respectivamente.

Nótese que al existir raíces unitarias en los polinomios característicos, la varianza no condicionada no existe (o es infinita). De hecho, es el mismo fenómeno que se encuentra cuando se consideran distribuciones incondicionales estables para series financieras.

Tomando como punto de partida un modelo GARCH(p,q) caracterizado por

$$E_{t-1}[\epsilon_t] = 0$$

$$E_{t-1}[\epsilon_t^2] = h_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}^2$$

con $\omega > 0$, $\alpha_i > 0$, $\beta_i > 0$. Si además,

$$1 - \sum_{i=1}^q \alpha_i \lambda^i - \sum_{i=1}^p \beta_i \lambda^i = 0$$

Entonces existen $d > 0$ raíces unitarias y $(\max(p,q)-d)$ raíces fuera del círculo unitario. El modelo se denomina integrado en varianza de orden d si $\omega = 0$, e integrado en varianza de orden d con deriva si $\omega > 0$.

El modelo IGARCH puede parecer similar en términos de predicción al de camino aleatorio para la media condicional. En el modelo de camino aleatorio con deriva la predicción de la media s periodos en el futuro es igual a la de hoy y los shocks en el sistema son permanentes. Del mismo modo, los shocks de la varianza condicionada cuando se manifiesta la existencia del IGARCH son también permanentes.

Pero a diferencia del camino aleatorio, o más genéricamente de los modelos ARIMA, el modelo IGARCH es estrictamente estacionario y ergódico (aunque no estacionario en covarianza). Además, mientras la teoría asintótica para los modelos ARCH es extremadamente difícil, **Lumsdaine (1991)** demuestra que los procedimientos estándar de inferencia asintótica son generalmente válidos incluso cuando los datos son generados por un modelo IGARCH, si bien las evidencias de Monte Carlo presentadas en **Hong (1988)** sugieren que los tamaños muestrales deben ser bastante grandes. Por otra parte, **Nelson (1990b)** demostró que un (I)GARCH con media nula, a diferencia de los modelos ARIMA, es empíricamente imposible, pues converge a cero con rapidez.

Muchos de los modelos GARCH empíricos estimados para series financieras son integrados en varianza. Según **Nelson (1992)** este fenómeno no es más que una consecuencia de las altas frecuencias de observación de los datos. En este sentido, cuando estas mismas series financieras son observadas a frecuencias bajas la evidencia de una raíz unitaria desaparece. **Lamoreaux y Lastrapes (1990)** sugieren que la aparente integración en varianza es el resultado de cambios estructurales.

La existencia de raíces unitarias en las varianzas se puede contrastar de manera análoga al caso de las medias aunque, a diferencia de las raíces unitarias en media, los contrastes se distribuyen de una forma estándar (**Engle y Bollerslev, 1986**). Del mismo modo, y en analogía a la cointegración en media, la persistencia en varianza puede ser común entre diferentes series. En este sentido, las variables se dice que presentan co-persistencia o cointegración en varianza (**Engle y Bollerslev, 1986; Kunst, R.M., 1992**), lo cual tiene importantes implicaciones en la construcción del predictor óptimo a largo plazo para las varianzas y covarianzas condicionadas.

Modelo ARCH-M

El modelo ARCH-M (**Engle, Lilien y Robins, 1987**) pretende modelizar, simultáneamente, la evolución de la media y varianza de la serie temporal.

$$y_t = g(x_{t-1}, h_t^2; \beta) + \varepsilon_t$$

La forma funcional de $g()$ suele ser lineal o logarítmica. En este modelo un incremento de la varianza condicionada estará asociado con un incremento o decremento de la media condicionada de y_t dependiendo del signo de la derivada parcial de $g(x_{t-1}, h_t^2, \beta)$ con respecto a h_t^2 .

Este tipo de modelos puede ser muy apropiado para representar muchas teorías financieras en las que se supone que la rentabilidad depende del riesgo, medido por la varianza condicionada.

Una extensión del ARCH-M es el ARCH-M de parámetros cambiantes, TVP-ARCH-M, (Chou, Engle y Kane, 1992).

$$\begin{aligned} y_t &= b_t h_t^2 + e_t \\ b_t &= b_{t-1} + v_t \\ h_t^2 &= \alpha_0 + \alpha_1 \eta_{t-1}^2 + \alpha_2 h_{t-1}^2 \end{aligned}$$

donde, e_t y v_t son i.i.d. $N(0, h_t)$ y $N(0, \sigma^2)$, respectivamente, y están incorrelacionados. η_t aparece porque b_t y h_t^2 son inobservables y no es más que,

$$\eta_t = y_t - E_{t-1}(y_t) = y_t - E_{t-1}(b_t) h_t^2 = e_t + [b_t - E_{t-1}(b_t)] h_t^2$$

$E_{t-1}(b_t)$ es el predictor óptimo de b_t en $t-1$. Nótese que cuando h_t^2 es observable las ecuaciones para y_t y b_t no son más que el modelo de regresión de parámetros cambiantes.

Modelos Vectoriales o Multivariantes

Muchos de los problemas teóricos financieros no pueden ser analizados de forma univariante. En este sentido, los modelos vectoriales o multivariantes se han desarrollado para explicar la variación simultánea de las varianzas heterocedásticas de diversas variables.

En el modelo Multivariante o ARCH Vectorial existe una representación para ϵ_t igual a un vector $N \times 1$ de procesos estocásticos. Así,

$$\epsilon_t = z_t \Omega^{-1/2}_t \text{ con } z_t, \text{ i.i.d., } E(z_t) = 0 \text{ y } \text{VAR}(z_t) = I.$$

Ω_t es una matriz $N \times N$ de varianzas-covarianzas semidefinida positiva y medible con respecto al conjunto de información en el período $t-1$.

En el modelo Multivariante ARCH(q) (Kraft y Engle, 1983) Ω_t está definida por una función lineal entre los productos cruzados de los errores pasados contemporáneos.

$$\text{vech}(\Omega_t) = \omega + \sum_{i=1}^q A_i \text{vech}(\varepsilon_{t-i} \varepsilon'_{t-i})$$

donde $\text{vech}(\varepsilon_{t-i} \varepsilon'_{t-i})$ es el operador que permite transformar la submatriz diagonal inferior de una matriz $N \times M$ en un vector $(N(N+1)/2) \times 1$, ω es un vector $(N(N+1)/2) \times 1$, y A_i es una matriz de orden $(N(N+1)/2) \times (N(N+1)/2)$.

El modelo ARCH multivariante puede ampliarse a un modelo GARCH(p,q) Multivariante (Bollerslev, Engle and Wooldridge, 1988) del tipo

$$\text{vech}(\Omega_t) = \omega + \sum_{i=1}^q A_i \text{vech}(\varepsilon_{t-i} \varepsilon'_{t-i}) + \sum_{i=1}^p B_i \text{vech}(\Omega_{t-i})$$

que introduce un sumatorio con los términos en $\text{vech}(\Omega_{t-i})$ y un término B_i que es una matriz de orden $(N(N+1)/2) \times (N(N+1)/2)$.

Obsérvese que en este modelo hay $1/2N(N+1)[1+N(N+1)(p+q)/2]$ parámetros, por lo que incluso para N pequeños deben imponerse restricciones. En este sentido se han desarrollado los ARCH Factoriales (Engle (1987), Engle, NG y Rothschild (1990), Diebold y Nerlove (1989)) en los que se incluyen factores latentes no observables. Estos factores pueden estar motivados por una aparente comunalidad en las agrupaciones de las volatilidades entre las diferentes variables.

En estos modelos la matriz de varianzas condicionadas Ω_t está expresado por una combinación lineal de $\varepsilon_{t-i} \varepsilon'_{t-i}$ y Ω_{t-i}

$$\Omega_t = v + \sum_{i=1}^q \sum_{k=1}^K a_{ik}^2 g_k f_k' \varepsilon_{t-i} \varepsilon'_{t-i} f_k g_k' + \sum_{i=1}^p \sum_{k=1}^K b_{ik}^2 g_k f_k' \Omega_{t-i} f_k g_k'$$

a_{ik} , b_{ik} son escalares constantes, g_k y f_k son vectores $N \times 1$ con las propiedades de que

$$f_k' g_k = 1 \text{ y } f_k g_j = 0 \text{ para } k \text{ distinto de } j.$$

Así, la representación GARCH para los k factores sugerida por Engle (1987)

$$\text{vech}(\Omega_t) = W \sum_{k=1}^K \text{vech}(g_k g_k') + \sum_{k=1}^K \text{vech}(f_k f_k') h_{kt}^2$$

donde, $W = \text{vech}(v)$ y h_{kt}^2

$$h_{kt}^2 = w_k + \sum_{i=1}^q a_{ik}^2 (f_k' \varepsilon_{t-i})^2 + \sum_{i=1}^p b_{ik}^2 h_{kt-i}^2$$

Modelo Estructural de Series Temporales con errores ARCH (STARCH)

Este modelo propuesto por Harvey, Ruiz y Sentana (1992) parte de la idea de que los errores que se obtienen del modelo estructural de series temporales básico, BSM, (Harvey, 1989) tienen una conducta heteroscedástica y modelizable mediante la representación ARCH.

Este modelo, como los modelos estructurales de series temporales univariantes, puede ser expresado en forma de espacio de estados. La ecuación de observación o medida es,

$$y_t = Z_t \alpha_t + X_t \beta + \Lambda \varepsilon_t + \varepsilon_t$$

donde, α_t es un vector de estado $m \times 1$; modelizable según la siguiente ecuación de transición. Ésta determina al vector de estado implícito en la ecuación anterior,

$$\alpha_t = T_t \alpha_{t-1} + \psi \eta_t + \eta_t$$

Z_t y T_t son matrices $N \times M$ y $M \times M$, respectivamente, que pueden ser no estocásticas, X_t es una matriz $N \times K$ de variables observables, β es un vector $K \times 1$, ε_t y η_t , que se distribuyen $NID(0, H^*)$ y $NID(0, Q^*)$, respectivamente, y están incorrelacionados.

El efecto ARCH se introduce como un modelo lineal para las perturbaciones ε_t y η_t . Las ecuaciones para el modelo de varianza heterocedástica se concretan en,

$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

para el error de la primera ecuación del modelo estructural, y,

$$q_t = \gamma_0 + \gamma_1 \eta_{t-1}^2$$

para el modelo de la ecuación de estado. Evidentemente, ϵ_t , ϵ_t^* , η_t y η_t^* son mutuamente independientes. Si se eliminase ϵ_t y η_t se tiene el modelo lineal de espacios de estados estándar.

Las perturbaciones ϵ_t están serialmente incorrelacionadas, poseen media nula y varianza $\alpha_0/(1-\alpha_1)$ con $\alpha_0 > 0$ y $0 < \alpha_1 < 1$. Además, las perturbaciones

$$\Lambda \epsilon_t + \epsilon_t^*$$

son ruido blanco multivariante con matriz de covarianzas

$$\frac{\alpha_0}{(1-\alpha_1)} \Lambda \Lambda' + H'$$

Además, η_t está serialmente incorrelacionado, con media nula y varianza $\gamma_0/(1-\gamma_1)$ con $\gamma_0 > 0$ y $0 < \gamma_1 < 1$, y con

$$\psi \eta_t + \eta_t^*$$

que también es ruido blanco multivariante.

Por ejemplo, en el modelo estructural más simple donde

$$y_t = \mu_t + \epsilon_t$$

$$\mu_t = \mu_{t-1} + \eta_t$$

donde y_t es la serie temporal, ϵ_t y η_t son ruido blanco y ecuación de transición

$$\alpha_t = \begin{bmatrix} \mu_t \\ \eta_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \eta_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \eta_t$$

3.2.- MODELOS NO LINEALES PARA LA VARIANZA HETEROSCEDÁSTICA.

Hipótesis de la Forma Funcional

Una segunda línea de investigación sobre los modelos ARCH considera la forma funcional de la varianza condicionada, h_t^2 , distinta de la especificación lineal inicialmente propuesta. Engle (1982), Engle y Bollerslev (1986), Geweke (1986) y Pantula (1986) han sugerido otras formas funcionales alternativas que intentan especificar una mejor conducta de los movimientos de la varianza condicionada en el tiempo.

Higgins y Bera (1990) señalan que la especificación de una función correcta para la varianza condicionada resulta bastante importante, porque la precisión de los intervalos de confianza depende de la selección correcta de una función que relacione las varianzas futuras con el conjunto de información, y, porque la precisión de los contrastes que se utilizan para detectar la existencia de procesos de varianza heterocedástica se determina parcialmente por la forma funcional que se elija para la misma. Así, una forma funcional incorrecta para los errores obtenidos de la estimación de cualquier modelo de regresión, puede conducir a resultados inconsistentes en las estimaciones de los parámetros. Estos modelos no lineales pueden resumirse en dos grandes grupos según el tipo de estimación realizada, estimación paramétrica y estimación no paramétrica.

1.- MODELOS NO LINEALES Y ESTIMACIÓN PARAMÉTRICA

Estos modelos se caracterizan porque en los procesos de estimación de los parámetros del modelo se utilizan distribuciones de densidad condicionadas paramétricas (normal condicionada, t-Student, etc.).

Modelos ARCH exponenciales y de valor absoluto.

Los modelos ARCH no lineales ya habían sido propuestos por Engle (1982, pp 993). Éste propone el siguiente modelo exponencial:

$$h_t^2 = \exp (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2)$$

en el que la varianza es positiva para todos los valores de α . Sin embargo los datos generados por este tipo de proceso pueden implicar una varianza infinita dificultando la estimación y la inferencia.

Como alternativa se ha propuesto el modelo de valor absoluto,

$$h_t^2 = \alpha_0 + \alpha_1 |\varepsilon_{t-1}| + \dots + \alpha_p |\varepsilon_{t-p}|$$

en el que los parámetros de la varianza, α , son positivos, $\alpha > 0$, aunque a diferencia del modelo anterior la varianza es finita.

Modelo Logarítmico (log-ARCH)

Adicionalmente, y en un intento de resolver los problemas ocasionados por el incumplimiento de la hipótesis de no-negatividad en el modelo ARCH, Geweke (1986) y Pantula (1986) sugieren una forma funcional alternativa a las propuestas por Engle (1982), asegurando una varianza condicionada positiva y finita para todo el espacio paramétrico de α . Formalmente,

$$\log (h_t^2) = \alpha_0 + \alpha_1 \log (\varepsilon_{t-1}^2) + \dots + \alpha_p \log (\varepsilon_{t-p}^2)$$

Modelo ARCH Exponencial Generalizado (EGARCH)

Otra forma funcional que está dentro de las especificaciones no lineales es el modelo Exponencial ARCH Generalizado de orden (p,q), o EGARCH (p,q), propuesto por Nelson (1990c),

$$\log h_t^2 = \alpha + \frac{(1 + \psi_1 L + \dots + \psi_q L^q)}{(1 - \Delta_1 L - \dots - \Delta_p L^p)} g(z_{t-1})$$

en el que

$$g(z_t) \equiv \theta z_t + \gamma[|z_t| - E|z_t|]$$

sustituyendo en la ecuación en logaritmos

$$\log h_t^2 = \omega + \sum_{i=1}^q \alpha_i [\theta z_{t-i} + \gamma[|z_{t-i}| - E|z_{t-i}|]] + \sum_{i=1}^p \beta_i \log h_{t-i}^2$$

De forma alternativa (Geweke, 1986 y Pantula, 1986)

$$\log (h_t^2) = \omega + \sum_{i=1}^q \alpha_i \log (z_{t-i}^2) + \sum_{i=1}^p \beta_i [\log (z_{t-i}^2) - \log (h_{t-i}^2)]$$

A diferencia del GARCH(p,q) este modelo no necesita restricciones en los parámetros para asegurar la no negatividad de la varianza condicionada. Por otra parte, obsérvese que en este modelo h_t^2 es una función asimétrica de los valores pasados de ϵ_t . Es decir, la varianza no sólo depende de la magnitud sino también del signo de los shocks. De hecho este fenómeno de 'apalancamiento' se puede observar con cierta frecuencia en el comportamiento de los mercados de acciones.

Modelo no Lineal ARCH (NARCH)

Un modelo que contiene muchos de los propuestos en la literatura es el desarrollado por **Higgins y Bera (1990)**. El planteamiento de su modelo se centra en el desarrollo para la varianza condicionada de una función similar a la que se conoce por la teoría económica como función CES o de Elasticidad Constante de Sustitución, que comprende, como casos particulares, a las funciones de producción más conocidas como la función de Leontieff, la función de Cobb-Douglas o funciones de tipo lineal,

$$h_t^2 = [\phi_0 (\sigma^2)^\delta + \phi_1 (\epsilon_{t-1}^2)^\delta + \dots + \phi_p (\epsilon_{t-p}^2)^\delta]^{1/\delta}$$

$$\sigma^2 > 0, \phi_i > 0 \text{ y } 0 < \delta < 1 \text{ para } i=0,1,\dots,p, \text{ y } \sum \phi_i = 1$$

El NARCH(p) es una especificación bastante flexible, y puede reescribirse como,

$$\frac{h_t^{2\delta} - 1}{\delta} = \phi_0 \frac{(\sigma^2)^\delta - 1}{\delta} + \phi_1 \frac{(\epsilon_{t-1}^2)^\delta - 1}{\delta} + \dots + \phi_p \frac{(\epsilon_{t-p}^2)^\delta - 1}{\delta}$$

que no es más que una transformación tipo **Box-Cox (1964)**. Así cuando $\delta=1$ el modelo para la varianza condicionada deviene lineal y cuando $\delta \rightarrow 0$ el modelo es logarítmico,

$$\log(h_t^2) = \phi_0 + \phi_1 \log(\epsilon_{t-1}^2) + \dots + \phi_p \log(\epsilon_{t-p}^2)$$

La varianza del modelo NARCH(1) es finita siempre que,

$$\phi_1 (\pi^{-\frac{1}{2}}) (2^\delta) \Gamma[\frac{\delta+1}{2}] < 1$$

donde $\Gamma(\cdot)$ es la función Gamma.

El modelo propuesto por **Higgins y Bera (1990)** fue propuesto para predecir el comportamiento futuro de varios tipos de cambio con respecto al dólar. Si bien las predicciones puntuales obtenidas por este modelo son similares a las obtenidas por otros modelos más sencillos, las predicciones por intervalo son mucho más precisas en el caso del NARCH, por lo que parece que este modelo capture más óptimamente los movimientos en el tiempo de la varianza condicionada.

2.- MODELOS NO LINEALES Y ESTIMACION NO PARAMETRICA.

En estos modelos la estimación de los parámetros se realiza con funciones de densidad no paramétricas, generalmente aproximaciones polinomiales y expansiones trigonométricas, que evitan la utilización de la hipótesis de normalidad condicionada para la distribución de los residuos del modelo de regresión o serie temporal.

Pagan y Schwert (1990), basándose en **Gallant (1981)**, analizan una alternativa no paramétrica en la cual especifican la varianza condicionada utilizando un desarrollo en serie.

$$h_t^2 = \sigma^2 + \sum_{j=1}^L [(\alpha_j z_{tj} + \beta_j z_{tj}^2) + \sum_{k=1}^2 (\gamma_{jk} \cos (kz_{tj}) + \delta_{jk} \sin (kz_{tj}))]$$

h_t^2 , se aproxima por una función polinomial y por términos trigonométricos de los errores estimados, $z_{tj} = \epsilon_{t-j}$.

Como alternativa **Pagan y Schwert (1990)** utilizan también funciones de kernel para aproximar el proceso lineal discreto que modeliza la varianza condicionada. En concreto

$$h_t^2 = \sum_{j=1}^T \omega_{jt} \epsilon_j^2 \quad ; \quad \sum_{j=1}^T \omega_{jt} = 1$$

en esta expresión ω_{jt} son ponderaciones construidas recurriendo a funciones de kernel,

$$\omega_{jt} = \frac{K(z_t - z_j)}{\sum_{k=1}^T K(z_k - z_t)}$$

K son funciones de Kernel caracterizadas por ser no nulas, simétricas e integrables.

Hipótesis de Normalidad Condicionada

Como vimos, la distribución condicionada de las observaciones generadas a partir de los modelos ARCH es frecuentemente asimétrica y leptocúrtica (es decir, más apuntada que la distribución normal teórica). Sin embargo, aunque es frecuente que los residuos estandarizados de los modelos estimados, $z_t = \epsilon_t / h_t$, tengan elevados grados de leptocurtosis, si la varianza está correctamente especificada los excesos de curtosis no pueden exceder a los de los residuos sin estandarizar (Hsieh, 1989).

Para eliminar el elevado grado de leptocurtosis de las distribuciones empíricas se han propuesto, alternativamente, distribuciones condicionadas con colas más gruesas que la distribución normal: la distribución de densidad t-Student (Bollerslev, 1987), la distribución normal-log-normal en Hsieh (1989), o la exponencial generalizada en Nelson (1990c), entre otros. También existen aproximaciones de distribuciones de densidad semiparamétricas en las que se utilizan expansiones polinomiales (Bollerslev, Chou y Kroner, 1992).

4.- ESTIMACIÓN.

Los parámetros de un modelo ARCH se pueden estimar por el método de Máxima Verosimilitud (Engle (1982)), el Método Generalizado de los Momentos (MGM) (Mark (1988)), o por métodos de inferencia bayesiana (Geweke (1989)). Particularmente, el estimador utilizado por Engle (1982) es el Máximo Verosímil, que es asintóticamente superior a los MCO, los cuales no alcanzan la cota de Cramer-Rao.

En los modelos para la varianza heterocedástica condicionada el logaritmo de la función de verosimilitud se construye bajo el supuesto de normalidad condicionada del error. Sea $L(\alpha)$ es el logaritmo de la función de verosimilitud para el vector de parámetros α y todos los valores de ϵ^2 , $(\epsilon^2_1, \dots, \epsilon^2_T)$; l_t el logaritmo de la función de verosimilitud para la t-ésima observación, entonces

$$L(\alpha) = \sum_{t=1}^T \frac{l_t(\alpha)}{T} = -\frac{1}{2T} \sum_{t=1}^T \left[\log h_t^2 + \frac{\epsilon_t^2}{h_t^2} \right]$$

ϵ_i/h_i es z_i , los residuos estandarizados.

Para estimar α se maximiza dicha expresión. Las condiciones de primer orden son,

$$\frac{\partial J_i}{\partial \alpha} = \frac{1}{2h_i^2} \frac{\partial h_i^2}{\partial \alpha} \left[\frac{\epsilon_i^2}{h_i^2} - 1 \right]$$

y la segunda derivada,

$$\frac{\partial^2 J_i}{\partial \alpha \partial \alpha'} = \frac{-1}{2h_i^4} \frac{\partial h_i^2}{\partial \alpha} \frac{\partial h_i^2}{\partial \alpha'} \frac{\epsilon_i^2}{h_i^2} + \left[\frac{\epsilon_i^2}{h_i^2} - 1 \right] \frac{\partial}{\partial \alpha'} \left[\frac{1}{2h_i^2} \frac{\partial h_i^2}{\partial \alpha} \right]$$

En general los parámetros se suelen estimar eficientemente recurriendo a algún algoritmo iterativo tipo scoring. La i -ésima iteración de este algoritmo puede escribirse como,

$$\alpha^{(i+1)} = \alpha^{(i)} + (Z'Z)^{-1} Z'f^{(i)}$$

donde,

$$Z_i = (1, \epsilon_{i-1}^2, \dots, \epsilon_{i-p}^2) / h_i^{2(i)}$$

$$Z' = (Z_1', \dots, Z_n')$$

y f es el cociente entre la diferencia entre los errores ϵ_i^2 y h_i^2 y los valores de h_i^2 .

$$f_i^{(i)} = (\epsilon_i^2 - h_i^{2(i)}) / h_i^{2(i)}$$

$$f^{(i)} = (f_1^{(i)}, \dots, f_T^{(i)})$$

donde i expresa el número de iteraciones hasta que el parámetro maximice la función de verosimilitud. La ventaja de este algoritmo es que sólo necesita para su cálculo de las primeras derivadas de la función de verosimilitud, establecidas por las condiciones de primer orden.

Diversos autores señalan problemas de convergencia en la estimación máximo verosímil utilizando este tipo de algoritmos. Alternativamente, se han propuesto o bien utilizar otros métodos de estimación (**Mark (1988) Geweke, (1989)**) o bien otros algoritmos de optimización numérica (**Novales, 1993**).

La estimación de los parámetros de un GARCH(p,q) se realiza también por máxima verosimilitud, con la misma forma para el logaritmo de la función de verosimilitud,

$$L_t(\theta) = \sum_{i=1}^t \frac{l_i(\theta)}{T} = -\frac{1}{2} \sum_{i=1}^t \log(h_i^2) - \frac{1}{2} \sum_{i=1}^t \left(\frac{\varepsilon_i^2}{h_i^2} \right)$$

donde,

$$\begin{aligned} \frac{\partial l_t}{\partial \omega} &= \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \omega} \left(\frac{\varepsilon_t^2}{h_t^2} - 1 \right) \\ \frac{\partial h_t^2}{\partial \omega} &= z_t + \sum_{i=1}^p \beta_i \frac{\partial h_{t-i}^2}{\partial \omega} \end{aligned}$$

Se suele utilizar el algoritmo de **Berndt, Hall, Hall y Hausman, BHHH, (1974)**. Sea $\theta^{(i)}$ la estimación de los parámetros después de la estimación i -ésima, $\theta^{(i+1)}$ es calculado recursivamente como,

$$\theta^{(i+1)} = \theta^{(i)} + \lambda_i \left(\sum_{t=1}^T \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta^{(i)}} \right)^{-1} \sum_{t=1}^T \frac{\partial l_t}{\partial \theta}$$

siendo, λ_i una variable de longitud de paso que se aplica a cada iteración y que posibilita una rápida convergencia.

El mismo algoritmo se utiliza en los modelos ARCH multivariantes, en los que la función de verosimilitud que se trata de maximizar se expresa,

$$L(\theta) = \sum_{t=1}^T \left[\log f(\varepsilon_t, \Omega_t^{\frac{1}{2}}) - \log |\Omega_t^{\frac{1}{2}}| \right]$$

en la que $(-\log h_t)$ se ha sustituido por $(-\log |\Omega_t|)$, mientras que (ε_t/h_t) sustituye a (ε_t/Ω_t) .

El modelo STARCH se estima por máxima verosimilitud utilizando el Filtro de Kalman, al igual que en el modelo estructural básico propuesto por

Harvey (1989). En este caso, el filtro de Kalman añade a la estimación los parámetros del ARCH(1) para ϵ_t y η_t , es decir,

$$y_t = Z_t \alpha_t + X_t \beta + \Lambda \epsilon_t + \epsilon_t^* = [Z_t \ 0 \ \Lambda_t] \alpha_t^* + X_t \beta + \epsilon_t^*$$

con la ecuación de transición

$$\alpha_t^* = \begin{bmatrix} \alpha_t \\ \eta_t \\ \epsilon_t \end{bmatrix} = \begin{bmatrix} T_t & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_{t-1} \\ \eta_{t-1} \\ \epsilon_{t-1} \end{bmatrix} + \begin{bmatrix} I & \psi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_t^* \\ \eta_t \\ \epsilon_t \end{bmatrix}$$

A diferencia de los modelos anteriores la matriz de información obtenida bajo la hipótesis de normalidad condicionada de los modelos ARCH-M y EGARCH no es diagonal en bloques con respecto a los parámetros de la media y la varianza condicional. Esto significa que en estos modelos la estimación de los parámetros no puede realizarse independientemente. Así, la estimación consistente de los parámetros requiere la correcta especificación del modelo completo (para la media y para la varianza).

Bollerslev y Wooldridge (1991) proponen un método de estimación cuasi máximo verosímil en la que, en analogía a White, pretenden obtener una estimación consistente y robusta asintóticamente de la matriz de varianzas y covarianzas de los parámetros de un modelo ARCH.

El estimador cuasi máximo verosímil θ , se obtiene maximizando el logaritmo de la siguiente función de verosimilitud L_t ,

$$L_t(\theta) = \sum_{t=1}^T l_t(\theta) = -\frac{1}{2} \log |\Omega(\theta)| - \frac{1}{2} \epsilon_t(\theta)' \Omega^{-1} \epsilon_t(\theta)$$

Ω_t es la matriz de varianzas y covarianzas de las variables exógenas en un modelo de regresión $y_t = x_t \theta + \epsilon_t$, siendo y_t un vector de $k \times 1$ observaciones de la variable dependiente, x_t un vector de variables predeterminadas y ϵ_t un vector $k \times 1$ de residuos.

La matriz obtenida por **Bollerslev y Wooldridge (1991)** puede ser representada como

$$A_T^{-1} B_T A_T^{-1}$$

siendo las matrices A y B

$$A_T \equiv \frac{1}{T} \sum_{t=1}^T a_t(\theta_T)$$

$$B_T \equiv \frac{1}{T} \sum_{t=1}^T s_t(\theta_T)' s_t(\theta_T)$$

las cuales dependen de $E[y_t/x_t] = v_t(x_t, \theta)$, $V[y_t/x_t] = \Omega_t(x_t, \theta)$; $\epsilon_t(y_t, x_t, \Omega) = y_t - v_t(x_t, \theta)$ es un vector $k \times 1$ de residuos, y

$$\frac{\partial \mu_t(\theta)}{\partial \theta} = \nabla_{\theta} \mu_t(\theta)$$

$$\frac{\partial \Omega_t(\theta)}{\partial \theta} = \nabla_{\theta} \Omega_t(\theta)$$

son matrices de derivadas parciales de orden $K \times p$ y $K^2 \times p$, respectivamente, tal que,

$$a_t(\theta) = \nabla_{\theta} \mu_t(\theta)' \Omega_t^{-1}(\theta) \nabla_{\theta} \mu_t(\theta) + \frac{1}{2} \nabla_{\theta} \Omega_t(\theta)'$$

$$[\Omega_t^{-1}(\theta) \otimes \Omega_t^{-1}(\theta)] \nabla_{\theta} \Omega_t(\theta)$$

es la matriz de información condicionada, de orden $p \times p$, simétrica y positiva, y donde,

$$s_t(\theta)' = \nabla_{\theta} \mu_t(\theta)' \Omega_t^{-1}(\theta) \nabla_{\theta} \mu_t(\theta) + \frac{1}{2} \nabla_{\theta} \Omega_t(\theta)'$$

$$[\Omega_t^{-1}(\theta) \otimes \Omega_t^{-1}(\theta)] \nabla_{\theta} \Omega_t(\theta) \text{ vech } [\epsilon_t(\theta) \epsilon_t(\theta)' - \Omega_t(\theta)]$$

es un vector $1 \times p$.

Para funciones de distribución condicionales no normales se debe redefinir la función de verosimilitud. En este sentido, por ejemplo, se pueden utilizar las distribuciones del error generalizadas, GED, (Box y Tiao, 1975). Las GED no son una distribución sino una familia de distribuciones, incluyendo distribuciones más leptocúrticas y más platicúrticas que la distribución normal. En realidad la GED es una mezcla de distribuciones normales (Hsu, 1982). El logaritmo de la función de verosimilitud de una GED queda definido por,

$$L = t \text{ Lnd}(v) - \sum_{i=1}^T \left(\frac{1}{2} \text{Ln} h_i^2 + c(v) z_i^{\frac{2}{1+v}} \right) \quad -1 < v < \infty$$

$$z_i = \frac{\varepsilon_i}{h_i}$$

$$d(v) = \left[\Gamma\left(\frac{3(1+v)}{2}\right) \right]^{\frac{1}{2}} (1+v)^{-1} \left[\Gamma\left(\frac{1+v}{2}\right) \right]^{-\frac{3}{2}}$$

$$c(v) = \left[\Gamma\left(\frac{3(1+v)}{2}\right) \right]^{\frac{1}{1+v}} \left[\Gamma\left(\frac{1+v}{2}\right) \right]^{-\frac{1}{1+v}}$$

$\Gamma(\cdot)$ es la función gamma. Cuando $v=0$ z_i sigue una distribución normal estándar, y si $v>0$ ($v<0$) la distribución tiene colas más gruesas (delgadas) que la normal.

$v=1$, por ejemplo, se corresponde con la distribución doble exponencial.

5.- IDENTIFICACIÓN Y CONTRASTE DE UN MODELO ARCH

La existencia de un proceso ARCH(1) incrementa los errores estándar de las estimaciones de las autocorrelaciones muestrales. En concreto los errores estándar asintóticos serán

$$\frac{1+2\alpha_1}{\sqrt{T}}$$

mayores que los errores estándar de un proceso homocedástico ($T^{-1/2}$) (Weiss, 1984). Milhøj (1990) demuestra que esta diferencia provoca una pérdida en la potencia de los contrastes.

Así pues, cuando existe un esquema ARCH se debe obrar con cautela, a la hora de identificar a través de la FAS y FAP de ε_t^2 , puesto que al incrementarse los errores estándar las bandas utilizadas tradicionalmente no son las correctas.

Puesto que las herramientas tradicionales, FAS y FAP de ε_t^2 , no sirven para discernir la presencia y el orden de un modelo ARCH se han propuesto otros contrastes alternativos. La hipótesis nula que se trata de contrastar es la

existencia de homoscedasticidad frente a la existencia de heteroscedasticidad de orden p , o ARCH(p).

Los contrastes tradicionalmente utilizados para verificar la existencia de modelos con varianza condicionada están basados en los Multiplicadores de Lagrange (ML), que pueden expresarse de forma asintóticamente equivalente como

$$\xi = T f' z (z' z)^{-1} z' f / f' f = TR^2$$

donde R^2 es el coeficiente de determinación de la regresión de f sobre z . Como al añadir una constante y multiplicando por un escalar el R^2 de una regresión no se modifica, el R^2 anterior lo es también de la regresión de ϵ_t^2 sobre una constante y los p valores retardados de ϵ_t^2 .

Por ejemplo, en un ARCH(p)

$$h_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_p \epsilon_{t-p}^2$$

la hipótesis nula a contrastar es

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_p = 0$$

el estadístico a utilizar es,

$$TR^2 \sim \chi_p^2$$

siendo R^2 el coeficiente de determinación de la regresión. El estadístico se distribuye asintóticamente como una χ^2 con p grados de libertad, cuando la hipótesis nula es cierta.

La identificación del modelo GARCH tiene los mismos problemas asociados al ARCH. En este caso, la función de autocovarianzas es

$$\gamma_k = \sum_{i=1}^q \alpha_i \gamma_{k-i} + \sum_{i=1}^p \beta_i \gamma_{k-i} = \sum_{i=1}^m \varphi_i \gamma_{k-i}$$

donde, $k \geq p+1$, $m = \max(p, q)$, $\varphi_i = \alpha_i + \beta_i$ con $i = 1, \dots, q$. Así pues, la FAS es

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \sum_{i=1}^m \varphi_i \gamma_{k-i} ; \quad n \geq p+1$$

En el contraste del modelo GARCH(p,q) se plantea la hipótesis nula es la no existencia del GARCH contra la alternativa de que exista. Nótese que la hipótesis nula puede contener raíces del ARCH(p), es decir, $\alpha_i > 0$, para $i=1, \dots, p$.

Si se realiza una partición de la varianza condicionada, tal que,

$$h_t = Z_t' \omega = Z_1' \omega_1 + Z_2' \omega_2$$

correspondiendo ω_2 a la parte 'media móvil' del GARCH, entonces la hipótesis nula puede plantearse como

$$H_0 : \omega_2 = 0$$

es decir,

$$H_0 : (\alpha_p, \dots, \alpha_q, \beta_p, \dots, \beta_p) = (\alpha_p, \dots, \alpha_q, 0, \dots, 0)$$

Por tanto rechazar la hipótesis nula implica que se rechaza la existencia de que la varianza condicionada siga un proceso estocástico lineal discreto GARCH, pero no se puede asegurar que la varianza no sea heterocedástica, es decir la existencia de efectos ARCH.

Teniendo en cuenta esta limitación, el contraste de un modelo GARCH(p,q) está determinado por el siguiente estadístico definido a través de los Multiplicadores de Lagrange,

$$LM = \frac{1}{2} [f' Z (Z' Z)^{-1} Z' f]$$

donde f es,

$$f = \left[\frac{\varepsilon_1^2}{h_1^2} - 1, \dots, \frac{\varepsilon_T^2}{h_T^2} - 1 \right]$$

y Z se es un vector formado por las derivadas parciales tal que

$$Z = \left[h_1^2 \frac{\partial h_1^2}{\partial \omega}, \dots, h_r^2 \frac{\partial h_r^2}{\partial \omega} \right]$$

que se distribuye como

$$TR^2 \sim \chi_r^2$$

donde r es el número de elementos de ω_2 .

La existencia de asimetrías (modelos tipo EGARCH) puede ser investigada examinando el comportamiento de los residuos estandarizados de la estimación de modelos lineales. A este fin, **Engle y Ng (1991)** proponen tres sencillos contrastes. Con el contraste de 'sesgo de signo' se evalúan los efectos de las innovaciones positivas y negativas sobre la volatilidad. Sea S_{t-1}^- es una variable ficticia con valor igual a 1 cuando ϵ_{t-1} es negativo y 0 en otro caso, el estadístico de 'sesgo de signo' no es más que la t de Student en la siguiente regresión:

$$z_t^2 = a + bS_{t-1}^- + u_t$$

donde z_t son los residuos estandarizados. Los contrastes de 'sesgo de signo positivo (negativo)' evalúan el diferente impacto que las sorpresas grandes positivas (negativas) tienen sobre la volatilidad.

$$\begin{aligned} z_t^2 &= a + bS_{t-1}^-\epsilon_{t-1} + u_t \\ z_t^2 &= a + bS_{t-1}^+\epsilon_{t-1} + u_t \end{aligned}$$

Finalmente, se pueden contrastar conjuntamente los tres efectos a través de una F ,

$$z_t^2 = a + b_1S_{t-1}^- + b_2S_{t-1}^-\epsilon_{t-1} + b_3S_{t-1}^+\epsilon_{t-1} + u_t$$

BIBLIOGRAFÍA

- **Berndt, E.K., Hall, B.H., Hall, R.E. and Hausman, J.A. (1974):** "Estimation Inference in Nonlinear Structural Models". Annals of Economic and Social Measurement, 4, 653-665.

- Bera, A.K. and Higgins, M.L. (1992): "A Survey of ARCH Models: Properties, Estimation and Testing", University of Illinois at Urbana-Champaign
- Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity". Journal of Econometrics, 31, 307-327.
- Bollerslev, T. (1987): "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return". Review of Economics and Statistics, 69, 542-547.
- Bollerslev, T. and Wooldridge, J.M. (1991): "Quasi Maximum Likelihood Estimation and Inference in Dynamic Models with Time Covariances" Econometrics Reviews.
- Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992): "ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence". Journal of Econometrics, 52, 5-59.
- Bollerslev, T., Engle, R.F. and Wooldridge, J.M. (1988): "A Capital Asset Pricing Model with Time Varying Covariances", Journal of Political Economy, 96, 116-131.
- Box, G.E.P. and Cox, D.R. (1964): "An Analysis of Transformations". Journal of the Royal Statistical Society, Serie B, 26, 211-243.
- Box, G.E.P. and Jenkins, G.M. (1976): Time Series Analysis, Forecasting and Control, Holden Day, San Francisco.
- Box, G.E.P. and Tiao, G.C. (1975): Bayesian Inference in Statistical Analysis, Addison-Wesley, Reading, Ma.
- Brock, W.A., Hsieh, D. and LeBaron, B. (1991): Nonlinear Dynamics, Chaos and Instability, MIT Press, Cambridge.
- Chou, R., Engle, R.F. and Kane, A. (1992): "Measuring Risk Aversion from Excess Returns on a Stock Index". Journal of Econometrics, 52, 201-224.
- Diebold, F.X. and Nerlove, M. (1989): "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model". Journal of Applied Econometrics, 4, 1-21.

- Engle,R.F. (1982): "*Autoregressive Conditional Heteroskedasticity with Estimates of Variance of U.K. Inflation*". Econometrica, 50, 987-1007.
- Engle,R.F. (1987): "*Multivariate GARCH with factor structure. Cointegration in Variance* ", U.C.S.D. Department of Economics. Unpublished Manuscript.
- Engle,R.F. and Bollerslev,T. (1986): "*Modelling the Persistence of Conditional Variances*". Econometric Reviews, 5, pp 1-50, 81-87.
- Engle,R.F. and Ng V. (1991): "*Measuring and Testing the Impact of News on Volatility*" Discussion Paper, University of California, San Diego.
- Engle,R.F., Granger,C.W.J. and Kraft,D.F. (1984): "*Combining Competing Forecasts of Inflation using a Bivariate ARCH model*". Journal of Economics, Dynamics and Control, 8, 151-165.
- Engle,R.F., Ng V. and Rothschild,M. (1986): "*A Factor ARCH Model for Stock Returns*", Unpublished manuscript. Department of Economics, University of California, San Diego.
- Engle,R., Lilien,D.M.and Robins,R.P. (1987): "*Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model*". Econometrica, 55, 391-407.
- Gallant,A.R. (1981): "*On a Bias in Flexible Function Forms and a Essentially Unbiased Form*". Journal of Econometrics, 15, 211-244.
- Geweke,J. (1986): "*Modelling the Persistence of Conditional Variances: A Comment*". Econometric Reviews, 5, 57-61.
- Geweke,J. (1989): "*Exact Predictive Densities in Linear Models with ARCH Disturbances*". Journal of Econometrics, 44, 307-325.
- Granger,C.W.J., Robins,R.P.and Engle,R.F. (1984): "*Wholesale and Retail Prices: Bivariate Time Series Modelling with Forecastable Error Variances* ". en Belsley,D.A. and Kut,E. eds, Model Reliability, MIT Press, Cambridge, 1-17.
- Harvey,A.C. (1989): *Forecasting Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.

- Harvey,A., Ruiz,E. and Sentana,E. (1992): "*Unobserved Component Time Series Models with ARCH Disturbances*". Journal of Econometrics, 52, 129-157.
- Higgins,M.L. and Bera,A.K. (1990): "*A Class of Nonlinear ARCH Models*". Discussion Papers (Revised), University of California, San Diego, Department of Economics.
- Hong,C.H. (1988): "*The Integrated Generalized Autoregressive Conditional Heteroskedastic Model: The Process, Estimation and Monte Carlo Experiments*". Discussion Paper. Department of Economics, University of California, San Diego.
- Hsieh,D. (1989): "*Modelling Heteroskedastic in Daily Foreign Exchange Rates*". Journal and Business and Economic Statistics, 7, 307-317.
- Kraft,D.F. and Engle,R. (1983): "*Autoregressive Conditional Heteroskedasticity in Multiple Time Series*". Discussion Paper, Department of Economics, University of California.
- Kunst,R.M. (1992): "*Cointegrated ARCH Systems: Dinamic Patterns in Interest Rates*". Institute for Advanced Studies. Viena.
- Hsu,D. (1982): "*A Bayesian Robust Detection of Shift in the Risk Structure of Stock Market Returns*" Journal of the American Statistical Association, 77, 29-39.
- Lamoreaux,G.C. and Lastrapes,W.D. (1990): "*Persistence in Variance, Structural Change and the GARCH model*", Journal of Business and Economics Statistics, 8, 225-234.
- Lumsdaine, R.L. (1991): "*Asymptotic Properties of the Maximun Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models*". Discussion Paper, Department of Economics, Princeton.
- Mark.N. (1988): "*Time Varying Betas and Risk Premia in the Pricing of Forward Foreign Exchange Contracts*". Journal of Financial Economics, 22, 335-354.
- Milhøj,A. (1990): "*Distribution of Empirical Autocorrelations of Squared First Order ARCH Process*". Discussion Paper. Department of Statistics. University of Copenhagen.

- Nelson,D.B. (1990a) : "*ARCH Models as Diffusion Approximations*" Journal of Econometrics, 45, 7-38.
- Nelson,D.B. (1990b): "*Stationarity and Persistence in the GARCH(1,1) Model*" Econometric Theory, 6, 318-370.
- Nelson,D.B. (1990c): "Conditional Heteroskedasticity in Asset Returns: A New Approach". Econometrica, Vol.59, pp 347-370.
- Nelson,D.B. (1992): "*Filtering and Forecasting with Misspecified ARCH models: Getting the Right Variance with the Wrong Model*". Journal of Econometrics, 52, 61-90.
- Ng,V, Engle,R.F.and Rothschild,M. (1992) "*A Multi-Dynamic-Factor Model for Stock Return*". Journal of Econometrics, 52, 245-266.
- Nijman,Th.E. and Palm,F.C. (1991): "*Recent Developments in Modelling Volatility in Financial Data*", Center Discussion Paper 9168, Tilburg University.
- Novalés, A. and Gracia-Díez,M. (1993): "*Guía para la Estimación de Modelos ARCH*", Estadística Española, 35, No 132, 5-38.
- Pagan,A.R. and Schwert,G.W. (1990): "*Alternative Models for Conditional Stock Volatility*". Journal of Econometrics, 45, 267-290.
- Pantula,S.G. (1986): "*Modelling the Persistence of Conditional Variances: A Comment*". Econometric Reviews, 5, 71-73.
- Weiss,A.A. (1984): "*ARMA Models with ARCH Errors*". Journal of Time Series Analysis, 5, 129-143.
- Weiss,A.A. (1986): "*Asymptotic Theory for ARCH Models: Estimation and Testing*". Econometric Theory, 2, 107-131.