Estudios de Economía Aplicada
Nº7, 1997. Págs.101-129

Contrastes de especificación para los modelos de varianza Heterocedástica condicionada

PEREZ RODRIGUEZ, JORGE V.
Universidad de Barcelona
MURILLO FORT, CARLOS
Universidad Pompeu Fabra

Esta nueva versión incluye todas las correcciones sugeridas por el evaluador, las cuales nos han parecido oportunas y por las que les quedamos muy agradecidas.

RESUMEN

Partiendo de un esquema GARCH(p,q), presentamos una sinopsis de los contrastes estadísticos más relevantes que pueden ser aplicados a los modelos de varianza heterocedástica condicionada, y que pueden extenderse a todos tipo de especificaciones de la misma. Atendiendo a la distinción metodológica de Hendry sobre los contrastes de hipótesis en modelos de regresión, dividimos a éstos en contrastes de especificación (de modelos y de estabilidad de las varianzas) y en contrastes de mala especificación (de no linealidad, heterocedasticidad, curva de impacto de las noticias, capacidad predictiva y otros contrastes).

Palabras clave: ARCH, multiplicador de Lagrange, test de abarcamiento.

ABSTRACT

We show a sinopsis of the most relevant tests that are applied to the Autoregresive Conditional Heteroscedastic (ARCH) models, which can be extended to the class of these models. We use Hendry's methology distinction over hypothesis tests in regression models. In this way, we'll state a division in specification tests (models and variances stability) and misspecification tests (non linearity, heteroscedasticity, news impact curve, predict capability and others tests).

Key words: ARCH, Lagrange multiplier, Encompassing Test.

Artículo recibido en julio de 1996. Revisado en noviembre de 1996

1.- Introducción

La econometría tradicional ofrece múltiples formas de contrastar y seleccionar modelos. Ultimamente, el desarrollo de modelos serie temporales para la varianza heterocedástica ha proporcionado otro campo de aplicación de las diversas metodologías de contraste, no exentas de algunas complicaciones formales. Los distintos procedimientos de contraste de modelos econométricos permiten la obtención de la especificación adecuada y la evaluación desde el punto de vista econométrico de las teorías alternativas en la forma de métodos de comparación y selección de modelos. En la metodología de Hendry existen dos etapas en la selección de los mismos: los contrastes de mala especificación y los contrastes de globalización o especificación. Ambos se diferencian en la estructura de la hipótesis alternativa. En este sentido, la hipótesis alternativa es meramente instrumental en el caso de la mala especificación, es decir, rechazar la hipótesis nula no debe conducirnos a la aceptación de la hipótesis alternativa.

Desde el trabajo de Engle (1982) que introdujo la esencia del modelo ARCH más elemental, se han desarrollado una gran cantidad de manuscritos que investigaban las diversas parcelas y posibilidades que aquél trabajo original despertó. Este hecho se ha manifestado de diversas formas: especificaciones funcionales distintas para la varianza [Bollerslev (1986), Engle (1987), Engle, Kane y Noh (1993), Engle y Ng (1993), etc.], diversos métodos de estimación y funciones de verosimilitud a maximizar [Bollerslev (1986), Hsieh (1989), Bollerslev, Chou y Kroner (1992), etc.], propiedades serie temporales de los modelos estacionarios y no estacionarios o integrados [Weiss (1984 y 1986), Hong (1988), Nelson (1990b y 1992), Lumsdaine (1991 y 1995), etc.], contrastes lineales y no lineales de los modelos y de los parámetros [Chia-Shiang (1995), Lumsdaine (1995)].

Habiendo atendido las tres primeras consideraciones en otros trabajos [véase Pérez-Rodríguez (1995a,b)] intentaremos presentar una síntesis de algunos contrastes de hipótesis que pretenden encontrar la correcta especificación de la varianza condicionada heterocedástica (véase el cuadro 1 para un resumen), en la pretensión de recoger adecuadamente las regularidades empíricas que muestran muchas series económicas (como por ejemplo, el agrupamiento de la volatilidad, las distribuciones de probabilidad altamente leptocúrticas, los efectos apalancamiento o asimétricos y los efectos estacionales).

Cuadro 1.- Contrastes aplicados a la estimación de las varianzas condicionadas.

1) Contrastes de mala especificación:

a) Contrastes de no linealidad.
b) Contraste de heterocedasticidad.
c) Contrastes basados en la curva de impacto de las noticias.

* Sesgo de signo.

* Sesgo de signo negativo.

* Sesgo de signo positivo.

* Conjunto.

d) Contrastes de la capacidad predictiva dentro de la muestra.
e) Otros contrastes:

* Contraste de la varianza.

* Normalidad, asimetría y curtosis de los residuos estandarizados.

* Consistencia de la estimación de la varianza.

2) Contrastes de especificación:

2.1) Comparación entre modelos:

a) Modelos anidados: RV, Wald, LM.
b) Modelos no anidados: Encompassing test.

2.2) Contrastes de estabilidad de las varianzas.

El contenido de este trabajo se organiza como sigue. En la sección 2 se ilustra brevemente el modelo GARCH(p,q), que permitirá introducir algunos aspectos relevantes que utilizaremos a lo largo del trabajo. En la sección 3, ilustramos algunos

contrastes de mala especificación. Y, finalmente, en la sección 4, se desarrollan los contrastes de especificación, tanto sobre los modelos como sobre la estabilidad de los parámetros.

2.- El modelo GARCH(p,q)

La motivición de este tipo de modelos se justifica por la generación de varianzas que varían a lo largo del tiempo dentro de los modelos econométricos (por ejemplo, la consideración del riesgo en las decisiones de inversión, ahorro y producción de los agentes, y de esta forma, en la evaluación de la incertidumbre) y que son modelizadas como procesos estocásticos[1].

Supongamos que partimos de una especificación general de la varianza, la cual sigue un esquema GARCH(p,q) o ARCH generalizado de orden (p,q), donde el error obtenido proviene de un modelo AR(1), tal que

$$\begin{aligned} y_t &= \phi_0 y_{t-1} + \varepsilon_t \ , \ |\phi_0| < 1 \\ \varepsilon_t / \psi_{t-1} &\sim N\left(0, h_t^2\right) \\ h_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}^2 \end{aligned} \qquad [1]$$

siendo, y_t la serie temporal de interés; ε_t un error condicionado sobre el conjunto de información disponible en t-1 ($\psi_{\tau-1}$), que se distribuye normalmente con media nula y varianza condicionda heterocedástica, h^2_v modelizable como un proceso GARCH(p,q); y $\theta=[\alpha\ \beta]'$ es un vector de parámetros de orden (p+q+1)x1.

---

1 *Véase Bollerslev, Chou y Kroner (1992) para un excelente panorámica sobre la aplicación de este tipo de modelos al estudio de los mercados financieros, así como Holt y Aradhyula (1990) para una aplicación sencilla a la valoración de la volatilidad del precio esperado de la oferta agregada, entre otros.*

Cuadro 2.- Restricciones sobre los parámetros y sobre la varianza no condicionada.

Parámetros	ARCH(q)	GARCH(p,q)
θ	$\alpha_0 \geq 0$ $\alpha_i \geq 0$, i=1,..,q	$\alpha_0 \geq 0$ $\alpha_i \geq 0$, i=1,..,q $\beta_j \geq 0$, j=1,..,p
σ^2_ε	$\sum_{i=1}^{q} \alpha_i < 1$	$\sum_{i=1}^{q} {}_i + \sum_{j=1}^{P} \beta_j < 1$

Definiendo σ^2_ε como la varianza no condicionada, esta es igual a

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

Así, en el modelo GARCH(p,q) es suficiente que $\alpha_1+\ldots+\alpha_q+\beta_1+\ldots+\beta_p<1$, para que la varianza no condicionada sea estacionaria. De esta forma, cuando se inclumple esta condición, nos encontramos ante los modelos integrados en varianza, o IGARCH(p,q).

La estimación de los parámetros de un GARCH(p,q) se realiza por máxima verosimilitud, siendo el logaritmo de la función de verosimilitud de la expresión [1],

$$L_T\left(y_1,\ldots,y_T,h_0^2;\theta\right) = L_T(\theta) = \sum_{t=1}^{T} \frac{l_t(\theta)}{T} = -\frac{1}{2}\sum_{t=1}^{T}\log(h_t^2) - \frac{1}{2}\sum_{t=1}^{T}\left(\frac{\varepsilon_t^2}{h_t^2}\right) \qquad [2]$$

donde,

$$\frac{\partial l_t}{\partial \theta} = \frac{1}{2h_t^2}\frac{\partial h_t}{\partial \theta}\left(\frac{\varepsilon_t^2}{h_t^2} - 1\right)$$

$$\frac{\partial h_t^2}{\partial \theta} = z_t + \sum_{i=1}^{p}\beta_i \frac{\partial h_{t-i}^2}{\partial \theta}$$

Dado que las condiciones de primer orden son expresiones no lineales, se emplean algoritmos de optimización numérica para estimar el vector de parámetros. El algoritmo más utilizado es Berndt, Hall, Hall y Hausman (BHHH, 1974). Este método permite la estimación del vector de parámetros θ de forma iterativa (o recursiva) a partir de la siguiente expresión:

$$\theta^{(i+1)} = \theta^{(i)} + \lambda_i \left(\sum_{t=1}^{T}\frac{\partial l_t}{\partial \theta}\frac{\partial l_t}{\partial \theta'}\right)^{-1}\sum_{t=1}^{T}\frac{\partial l_t}{\partial \theta} \qquad [3]$$

donde, $\theta^{(i)}$ es la estimación de los parámetros después de la estimación i-ésima; y λ_i es una variable de longitud de paso aplicada en cada iteración y que posibilita una rápida convergencia para la obtención de $\theta^{(i+1)}$.

3.- Contrastes de mala especificación o especificación errónea

En la búsqueda de modelos congruentes o libres de errores de especificación, es decir, modelos en los que deben cumplirse una serie de supuestos mínimos que nos indican que el modelo está libre de error de especificación, la hipótesis alternativa es meramente instrumental, y por tanto rechazos de la hipótesis nula no implican aceptar la hipótesis alternativa.

En estos términos, la contrastación de especificaciones de tipo ARCH han derivado distintas expresiones que permiten validar si un modelo está correctamente especificado o no. Habitualmente, y la confirmación de la bondad de ajuste del modelo planteado, los contrastes aplicados se circunscribían a la utilización del multiplicador de Lagrange, obtenido de una regresión entre los scores de la función de verosimilitud en la especificación de la varianza [Engle (1982)]. Estos contrastes, sin embargo, no son los únicos. Con el paso del tiempo, se han postulado innumerables hipótesis que pretendían validar el modelo para la varianza heterocedástica estimado, y que veremos a continuación.

3.1.- Contrastes de no linealidad

Una estructura ARCH conduce a autocorrelación o dependencia lineal en la serie de los cuadrados de los residuos. Una de las opciones más utilizadas ha sido la identificación de la correlación serial en la serie de dichos cuadrados[2]. Para ello, se analiza la FAS de los errores al cuadrado, ε_t^2. En este caso, el coeficiente de correlación $\rho_2(\tau)$ de ε_t^2 es no nulo para órdenes de $\tau \geq 1$ (nótese que representamos los estadísticos con el subíndice 2 para identificar que se refieren al cuadrado de los errores, ε_t^2). Por ejemplo, en el caso de un ARCH(1) y GARCH(1,1) el coeficiente de autocorrelación de primer orden es igual a,

$$\rho_2(1) = \frac{2\alpha_0\alpha_1}{(1-\alpha_1)(1-3\alpha_1^2)} \neq 0$$

$$\rho_2(1) = \frac{\alpha_1(1-\alpha_1\beta_1-\beta_1^2)}{1-2\alpha_1\beta_1 \quad \beta_1^2} \neq 0$$

respectivamente. Además, las bandas de confianza sobre las correlaciones de los residuos cuando existen efectos ARCH no resultan válidas. En efecto, en este caso, las varianzas de las autocorrelaciones aumentan, viéndose modificadas al alza ante dicha presencia. Es por ello, que un criterio estadístico utilizado es el contraste de autocorrelación de Ljung-Box sobre los cuadrados ($Q_2(p)$), que nos permite valorar si las autocorrelaciones de los errores al cuadrado son ruido blanco o no lo son. Este contraste tiene gran potencia, aun cuando existan efectos ARCH. El estadístico se distribuye como χ_p^2, bajo la hipótesis nula de ruido blanco. Valores definidos en $Q_2(p)$ no significativos, es decir, $Q_2(p) < \chi_p^2$, implican que no existen estructuras dinámicas en los cuadrados.

Otro contraste que merece destacarse, ya que permite detectar patrones ocultos de comportamiento tales como las dependencias lineales y no lineales o las conductas no estacionarias, es el estadístico BDS [Brock, Decher y Sheinkman (1987)]. Este contraste está basado en la dimensión de las correlaciones temporales de los errores, y es un contraste de independencia.

3.2.- Contraste de homocedasticidad

El contraste de un modelo ARCH(q) como

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

implica contrastar la hipótesis nula: H_0: $\alpha_1=\ldots=\alpha_\theta=0$. Engle (1982) derivó un determinación de la regresión entre los errores del modelo estimado por mínimos cuadrados ordinarios, ε^2_t, con respecto a una constante y los q retardos del error, $\varepsilon^2_{t-1},\ldots,\varepsilon^2_{t-q}$ [aspecto que también argumentan Pesaran y Evans (1984)]. Bajo la hipótesis nula, TR^2 se distribuye como una χ^2_q, siendo q el número de restricciones. Por ejemplo, la regresión de ε_t^2 con respecto a sus retardos es

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

calculado, de este modo, el estadístico TR^2 (siendo T el número de observaciones empleadas en dicha regresión y R^2 el coeficiente de determinación de la regresión anterior) se distribuye como χ^2_1 bajo la hipótesis nula de homocedasticidad. Este estadístico permite probar la existencia de pertubaciones heterocedásticas.

En el contraste del modelo GARCH(p,q) definido en [1], se plantea como hipótesis nula la no existencia del GARCH contra la alternativa de que exista. Si se realiza una partición de la varianza condicionada, tal que,

$$h_t^2 = Z_t'\theta = Z_1'\theta_1 + Z_2'\theta_2$$

correspondiendo θ_2 a la parte media móvil del GARCH, entonces la hipótesis nula puede plantearse como $H_0 : \theta_2 = 0$, es decir,

$$H_0 : (\alpha_1,\ldots,\alpha_q,\beta_1,\ldots,\beta_p) = (\alpha_1,\ldots,\alpha_q,0,\ldots,0) \ .$$

Cuadro 3.- Multiplicador de Lagrange.

El valor del multiplicador de Lagrange puede obtenerse en términos del modelo estimado para la varianza condicionada heterocedástica a través de las puntaciones obtenidas en la función de verosimilitud:

$$LM = Tf'z(z'z)^{-1}z'f / f'f = TR^2$$

1) ARCH(1): La expresión asintóticamente equivalente puede escribirse comodonde R^2 es el coeficiente de determinación de la regresión de f sobre z. Donde f está definido por,

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

donde, z es un vector formado por las derivadas parciales primeras, tal que

$$z = \left[h_1^2 \frac{\partial h_1^2}{\partial \theta}, \ \ldots \ , h_T^2 \frac{\partial h_T^2}{\partial \theta}\right]$$

siendo, $\theta=[\alpha_0, \alpha_1, \ldots , \alpha_p]$.

2) GARCH(p,q). El multiplicador de Lagrange (LM) es

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

donde, $\theta=[\alpha_0, \alpha_1, \ldots , \alpha_q, \beta_1, \ldots , \beta_p]$. El estadístico se distribuye como $TR^2 \sim \chi_r^2$ donde r es el número de elementos de θ_2 y el R^2 es el coefciente de determinación de la regresión de f sobre z.

Por otro lado, el contraste de homocedasticidad frente a un proceso ARCH(q) o GARCH(p,q) es idéntico [Bollerslev (1986)].

Por tanto, rechazar la hipótesis nula implica el rechazo de la existencia de una varianza condicionada que sigue un proceso estocástico lineal discreto GARCH,

aunque no puede asegurarse que la varianza no sea heterocedástica, es decir que no existan efectos ARCH.

Nótese que la hipótesis nula puede contener raíces del ARCH(q), es decir, $\alpha_i > 0$, para $i = 1,\ldots,q$.

3.3.- Contrastes basados en la curva de impacto de las noticias

La curva del impacto de las noticias es una métrica que sirve para analizar el efecto de las noticias sobre la volatilidad condicionada heterocedástica, a través de la relación entre h^2_{t-1} y ε_{t-1}, y mide cómo la nueva información se incorpora dentro de la volatilidad estimada [véase Pagan y Schwert (1990)]. Precisamente, se la denomina curva de impacto de noticias porque hace referencia a ε_{t-1}, que son las noticias o shocks de la variable (positivos o negativos).

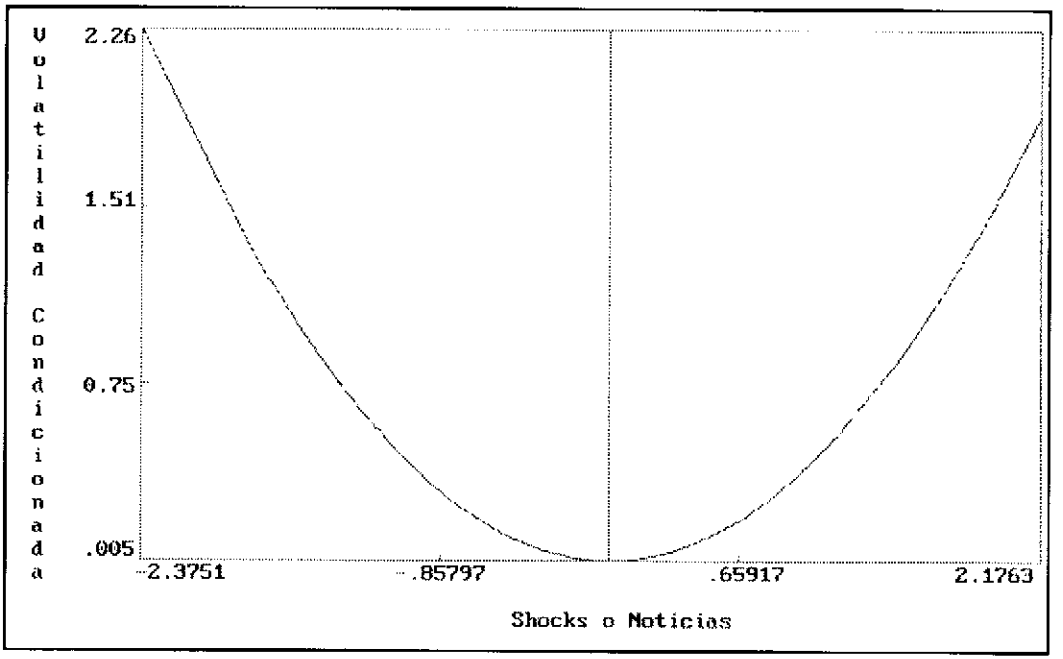
Cuadro 4.- Algunas especificaciones de la curva de impacto de las noticias.

Modelo	Especificación	Restricción
GARCH(1,1)	$h_t^2 = A + \alpha \varepsilon_{t-1}^2$	$\omega>0$, $0\leq\beta\leq1$, $0\leq\alpha\leq1$, $\sigma>0$, y $\alpha+\beta<1$
	$A = \omega + \beta\sigma^2$	
EGARCH(1,1)	$h_t^2 = A\exp\left[\frac{\gamma+\alpha}{\sigma}\varepsilon_{t-1}\right]$, $\forall \varepsilon_{t-1}>0$	$\omega>0$, $0\leq\beta\leq1$, $0\leq\alpha\leq1$, $\sigma>0$, $\alpha+\beta<1$, y $\gamma<0$
	$h_t^2 = A\exp\left[\frac{\gamma-\alpha}{\sigma}\varepsilon_{t-1}\right]$, $\forall \varepsilon_{t-1}\leq 0$	
	$A = \sigma^{2\beta}\exp\left[\omega - \alpha\sqrt{2\pi}\right]$	
GJR	$h_t^2 = A + \alpha\varepsilon_{t-1}^2$, $\varepsilon_{t-1}>0$	$\omega>0$, $0\leq\alpha\leq1$, $\sigma>0$, y $\gamma<0$
	$h_t^2 = A + (\alpha+\gamma)\varepsilon_{t-1}^2$, $\varepsilon_{t-1}\leq 0$	
	$A = \omega + \beta\sigma^2$	

De esta forma, la curva relaciona la varianza condicionada con el nivel de la varianza no condicionada ε_{t-1}^2 (Shocks o noticias de la variable retardada). Algunos ejemplos de funciones que modelizan la curva de impacto de las noticias, en el intento de medir como la nueva información es incorporada dentro de la volatilidad estimada, puede estudiarse para diversos modelos, a trvés de Engle y Ng (1993).

La curva de impacto de noticias está definida por una expresión forma cuadrática, siendo una curva simétrica y centrada en $\varepsilon_{t-1}=0$. Concretamente, σ es el parámetro de la varianza no condicionada (véase, por ejemplo el gráfico 1). El modelo EGARCH está representado por una función exponencial que posee varianzas más elevadas en los dos lados de la curva, que es asimetrica y centrada en $\varepsilon_{t-1}=0$.

Gráfico 1.- Curva del impacto de las noticias en un GARCH(1,1). Volatilidad condicionada (h_{t-1}) frente a los shocks o noticias (ε_{t-1}).



El modelo de Glosten, Jagannathan y Runkle (GJR) permite derivar una curva de impacto de las noticias que es asimétrica [véase el cuadro 4, y Engle y NG (1993)].

A partir de la métrica que acompaña a la curva de impacto de las noticias se definen distintas expresiones que permiten contrastar la especificación errónea del modelo de volatilidad. Estas son desarrolladas por Engle y Ng (1993), basándose en la idea de predecir los residuos cuadrados estandarizados, z^2_t, mediante alguna variable observada en el pasado que no está incluida en el modelo de volatilidad estudiado. Si estas variables pueden predecir los residuos normalizados, entonces el modelo estará mal especificado. De esta forma, estos autores proponen una batería de contrastes que

permiten examinar si el modelo está o no correctamente especificado, en base a la influencia que poseen los signos de los shocks o noticias.

El punto de partida es la especificación de un contraste que recibe el nombre de test de abarcamiento[2] (véase epígrafe 4.2).

De esta forma, Kraft y Engle (1983), Engle (1983) y Bollerslev (1986) consideran un contraste basado a probar que H_0: $\delta_a = 0$ en la ecuación [4]. Para ello, basta con realizar la siguiente regresión auxiliar, de tal forma que

$$z_t^2 = y_{0t}^* \delta_0 + y_{at}^* \delta_a + u_t$$

donde z_t son los residuos estandarizados ε_t/h_{0t}, y

$$y_{0t}^* = \frac{1}{h_{0t}^2}\frac{\partial h_t^2}{\partial \delta_0} \; ; \quad y_{at}^* = \frac{1}{h_{0t}^2}\frac{\partial h_t^2}{\partial \delta_a}$$

los cuales son evaluados en $\delta_a = 0$, δ_0 son los estimadores máximo verosímiles de δ, y u_t es un residuo.

Los contrastes ofrecidos por Engle y Ng (1993) se basan en cómo los shocks (es decir, valores de ε_{t-1}, positivos y negativos) que se producen en las variables afectan a la volatilidad, provocando que ésta sea mayor o menor dependiendo del signo de dichos shocks. Asumiendo, cierta pérdida en la potencia de los contraste, podemos eliminar arbitrariamente las variables y_{0t}^*, y únicamente regresar los valores de z_t^2 con respecto a los de y_{at}.

---

2 *Podemos considerlo como la versión española introducida por el profesor J.M. Otero de la voz anglosajona definida como Encompassing Test.*

Cuadro 5.- Distintos contrastes de sesgo de signo.

Contraste	Expresión	Decisión
Sesgo de signo	$z_t^2 = a + bS_{t-1} + u_t$	t(b=0)
Sesgo de signo (-)	$z_t^2 = a + bS_{t-1}^{-}\varepsilon_{t-1} + u_t$	t(b=0)
Sesgo de signo (+)	$z_t^2 = a + bS_{t-1}^{+}\varepsilon_{t-1} + u_t$	t(b=0)
Conjunto	$z_t^2 = a + b_1 S_{t-1}^{-} + b_2 S_{t-1}^{-}\varepsilon_{t-1} + b_3 S_{t-1}^{+}\varepsilon_{t-1} + u_t$	$F(b_1=b_2=b_3=0)$ o χ_3^2

. El contraste aplicado es un multiplicador de Lagrange igual a TR^2 que se distribuye como χ^2_m ,donde m es el número de restricciones (véase el cuadro 5, para obtener las distintas expresiones del contraste de sesgo de signo).

En este contraste es importante la inclusión de la variable S^{-}_{t-1}, construida de tal forma que

$$S_{t-1}^{-} = 1 \; ; \; si \; \varepsilon_{t-1} < 0$$

$$S_{t-1} = 0 \; ; \; si \; \varepsilon_{t-1} \geq 0$$

es decir, estab variable examina los impactos o shocks (positivos y negativos), sobre la volatilidad no predicha por el modelo. Contrastando la significación de los valores de b a través de una t-Student, el modelo planteado para la volatilidad heterocedástica es correcto cuando se acepta según el criterio t(b=0).

El contraste de sesgo de signo negativo utiliza en la ecuación de regresión la variable $S^{-}_{t-1}\ \varepsilon_{t-1}$. Se centra en los efectos que producen los diferentes shocks de valores negativos del término de error (grandes o pequeños) sobre la volatilidad que no es predecible por el modelo. Igualmente se trata de contrastar la significación del parámetro b (t(b=0)), donde si se acepta permite decidir que el modelo de varianza está correctamente especificado.

Por otro lado, el contraste de sesgo de signo positivo utiliza en la ecuación de regresión la variable $S^{+}_{t-1}\ \varepsilon_{t-1}$, siendo $S^{+}_{t-1} = 1 - S_{t-1}$. Se centra en los diferentes impactos que los shocks positivos (grandes y pequeños) tienen sobre la volatilidad que

no está explicada por el modelo. La decisión se toma estudiando el valor de t(b=0), y las conclusiones son idénticas a los casos anteriores.

Finalmente, el contraste conjunto de los tres efectos implica valorar la significación de los parámetros b_1, b_2 y b_3. Si estos parámetros son significativos conjuntamente, entonces el modelo para la varianza heterocedástica estará mal especificado. El contraste aplicable es un Multiplicador de Lagrange (ML), el cual es igual a TR^2, y se distribuye como χ_3^2, donde R^2 es el coeficiente de correlación múltiple de esta regresión, y 3 son los grados de libertad.

3.4.- Capacidad predictiva dentro de la muestra

La capacidad predictiva de los modelos puede calcularse recurriendo a diversas fórmulas elaboradas por Hamilton y Susmel (1994). Si la especificación del modelo planteado fuese correcta y los parámetros fuesen conocidos con certidumbre, entonces h^2_t será la esperanza condicionada de ε^2_t.

Cuadro 6.- Distintas funciones de pérdida.

	Estándar	Absoluto
Sobre errores	$MSE = \frac{\sum_{t=1}^{T} \left(\varepsilon_t^2 - h_t^2\right)^2}{T}$	$MAE = \frac{\sum_{t=1}^{T} \left\|\left(\varepsilon_t^2 - h_t^2\right)\right\|}{T}$
Sobre logaritmo errores	$LMSE = \frac{\sum_{t=1}^{T} \left(\ln(\varepsilon_t^2) - \ln(h_t^2)\right)^2}{T}$	$LMAE = \frac{\sum_{t=1}^{T} \left\|\ln(\varepsilon_t^2) - \ln(h_t^2)\right\|}{T}$

Este hecho permite definir diversas funciones de pérdida (véase el cuadro 6), como por ejemplo, el MSE es el error estándar medio; MAE es el error absoluto medio; LMSE es el error estándar del logaritmo de los errores; y LMSE es el error absoluto del logaritmo de los errores.

Sin embargo, debemos ser extramadamente cautos con la especificación MSE, por cuanto se basa en los momentos no condicionados de cuarto orden y estos no existen si los procesos son IGARCH. En cualquier caso, la buena capacidad predictiva implica que dichas medidas se encuentren próximas a cero.

3.5.- Otros contrastes

Finalmente, podemos considerar un grupo de criterios de validación del modelo estimado, que consideramos independientes y no incluibles en la taxonomías o categorías anteriores (véase el cuadro 7). En primer lugar, analizando que la varianza de los residuos al cuadrado sea mayor que la varianza de la varianza heterocedástica ($var(\varepsilon^2_t) \geq var(h^2_t)$), podríamos considerar que el modelo está correctamente especificado [Engle y Ng (1993)] por cuanto la varianza de la estimación es menor que la de los residuos al cuadrado obtenidos del modelo. Por otro lado, los residuos estandarizados ε_t/h_t deben aproximarse a una distribución normal estándar. Este aspecto es fundamental, si los residuos estandarizados no siguen la distribución normal parecería más conveniente realizar la estimación utilizando como función de verosimilitud una distribución que recogiese la forma y estructura de la posible distribución empírica de los datos. En este caso, resulta adecuado intentar proponer distintas funciones de verosimilitud como la t de Student, la función gamma, y en general, la familia distribuciones exponenciales generalizadas (GED) propuestas por Hsu (1982).

Adicionalmente, Engle (1982) advierte de la existencia de distribuciones más leptocúrticas para los residuos estandarizados. Concretamente, la asimetría y curtosis de los residuos estandarizados, ε_t/h_t, debería ser menor que la obtenida para los residuos sin estandarizar, ε_t [Hsieh (1989)], con lo cual, también el modelo para la varianza heterocedástica estaría correctamente especificado.

Cuadro 7.- Otras contrastes sobre la mala especificación.

Criterios	Decisión	Autores
$var(\varepsilon^2_t) \geq var(h^2_t)$	Correcta especificación	Engle y Ng (1993)
Normalidad de ε_t/h_t	Idem	Hsieh (1989)
Asimetría y curtosis de $\varepsilon_t/h_t < \varepsilon_t$	Idem	Hsieh (1989)
Contraste de $\varepsilon^2_t = \alpha + \beta h^2_t + v_t$	Consistencia e insesgadez de estimaciones de varianza si H_0: $\beta=1$ y $\alpha=0$	Pagan y Schwert (1990)

Y, finalmente, la consistencia de la estimación de la volatilidad puede analizarse mediante una regresión de los errores al cuadrado sobre una constante y las estimaciones de la varianza condicionada [Pagan y Schwert (1990)]. En este caso, la hipótesis nula que se contrasta es H_0: $\beta=1$, $\alpha=0$. Si ésta es aceptada, la especificación de la varianza es correcta.

4.- Contrastes de especificación

Una vez tenemos un conjunto de modelos rivales libres de errores de especificación se contrasta la especificación mediante una hipótesis nula y alternativa que tienen un carácter constructivo, es decir, la hipótesis alternativa está definida concretamente.

Este el caso de los contrastes que veremos a continuación sobre modelos de varianza heterocedástica y sobre la estabilidad de las varianzas.

4.1.- Contrastes sobre los modelos

En esta sección distinguiremos entre dos tipos de entornos: los modelos anidados y los modelos no anidados.

4.1.1.- Modelos anidados

Los contrastes utilizados cuando los modelos son anidados para restricciones lineales y no lineales son la razón de verosimilitud, el contraste de Wald y los multiplicadores de Lagrange. Partimos de un esquema serie temporal univariante, que se caracteriza porque la varianza del error es heterocedástica siguiendo un modelo GARCH(1,1) como [1], donde θ se obtiene maximizando la expresión [2]mediante [3].

Ahora, supongamos que θ_r es el vector de parámetros del modelo restringido, y θ_u es el vector de parámetros del modelo sin restringir. La comparación entre dos modelos mediante el contraste de la razón de verosimilitud (RV), que compara un modelo restringido definido en la hipótesis nula H_0: $\theta\in\Theta$, siendo Θ el espacio de parámetros de dicho modelo, y otro no restringido definido sobre H_a: $\theta\notin\Theta$, puede realizarse utilizando el estadístico de prueba $RV = -2(\text{Ln } H_0 - \text{Ln } H_a) \sim \chi^2_1$, bajo la hipótesis nula. Este contraste se basa en que si las restricciones que se contrastan fuesen ciertas, entonces no habría una reducción significativa en el valor máximo de la función de verosimilitud, por lo que la diferencia anterior, que es un escalar, debe ser pequeña. Así,

$$RV = -2T\left[L_T\left(y_1,\ldots,y_T,h_0^2;\hat{\theta}_r\right) - L_T\left(y_1,\ldots,y_T,h_0^2;\hat{\theta}_u\right)\right]$$

Otra forma de comparar los modelos econométricos es a través de los multiplicadores de Lagrange y el contraste de Wald. En el caso del multiplicador de Lagrange (LM) se trata de analizar el modelo no restringido. La expresión habitual está definida por

$$LM = \left(\frac{\partial L_T(\hat{\theta}_r)}{\partial \hat{\theta}_r}\right)' \left(I(\hat{\theta}_r)\right)^{-1} \left(\frac{\partial L_T(\hat{\theta}_r)}{\partial \hat{\theta}_r}\right) \sim \chi_q^2$$

siendo $L_T(\hat{\theta}_r)$ el logaritmo de la función de verosimilitud para el estimador restringido $\hat{\theta}_r$, I es la matriz de información, y q el número de restricciones bajo la hipótesis nula. Particularmente, los elementos que intervienen en la matriz de información pueden definirse como en Weiss (1986) donde, una estimación consistente de tales matrices

$$A_0 = -E\left[\frac{\partial^2 L_T(\theta_0)}{\partial\theta\,\partial\theta'}\right]$$

$$B_0 = E\left[T\frac{\partial L_T(\theta_0)}{\partial\theta}\frac{\partial L_T(\theta_0)}{\partial\theta'}\right]$$

puede realizarse en este caso a través de

$$A_T(\theta) = \frac{1}{2T}\sum_{t=1}^{T} h_t^{-4}\frac{\partial h_t^2}{\partial\theta}\frac{\partial h_t^2}{\partial\theta'} + \frac{1}{T}\sum_{t=1}^{T} h_t^{-2}\frac{\partial \varepsilon_t}{\partial\theta}\frac{\partial \varepsilon_t}{\partial\theta'}$$

$$B_T(\theta) = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial l_t}{\partial\theta}\frac{\partial l_t}{\partial\theta'}$$

El contraste de LM se basa en que si las restricciones que se contrastan son válidas, el estimador restringido debería estar próximo al punto que maximiza globalmente la funciónde verosimilitud, por lo que la pendiente de dicha función debería ser nula. Es un contraste basado en el valor numérico de dicha pendiente.

Atendiendo a su utilización, podemos distinguir entre dos formas de contraste Así, como argumento diferente a los explicados hasta el momento se introduce g(θ), que es una función de las restricciones[3].

Cuadro 8.- Estadísticos robustos y no robustos

Robusto	No Robusto
$LM_R = T\left[\left(\frac{\partial L_T}{\partial\theta}\Big\vert_{\hat\theta_r}\right)' A_T^{-1}(\hat\theta_r)\left(\frac{\partial g}{\partial\theta}\Big\vert_{\hat\theta_r}\right)\right] \times \left[\frac{\partial g}{\partial\theta'}\Big\vert_{\hat\theta_r} A_T^{-1}(\hat\theta_r) B_T(\hat\theta_r) A_T^{-1}(\hat\theta_r)\frac{\partial g}{\partial\theta}\right] \times \left[\frac{\partial g}{\partial\theta}\Big\vert_{\hat\theta_r} A_T^{-1}(\hat\theta_r)\left(\frac{\partial L_T}{\partial\theta}\Big\vert_{\hat\theta_r}\right)\right]$	$LM_{NR} = T\left[\left(\frac{\partial L_T}{\partial\theta}\Big\vert_{\hat\theta_r}\right)' A_T^{-1}(\hat\theta_r)\left(\frac{\partial L_T}{\partial\theta}\Big\vert_{\hat\theta_r}\right)\right]$
$W_R = T\, g(\hat\theta_u)'\left[\frac{\partial g}{\partial\theta'}\Big\vert_{\hat\theta_u} A_T^{-1}(\hat\theta_u) B_T(\hat\theta_u) A_T^{-1}(\hat\theta_u)\frac{\partial g}{\partial\theta}\Big\vert_{\hat\theta_u}\right]^{-1} g(\hat\theta_u)$	$W_{NR} = T\, g(\hat\theta_u)'\left[\frac{\partial g}{\partial\theta'}\Big\vert_{\hat\theta_u} A_T^{-1}(\hat\theta_u)\frac{\partial g}{\partial\theta}\Big\vert_{\hat\theta_u}\right]^{-1} g(\hat\theta_u)$

[3] *Por ejemplo, si se contrasta la hipótesis nula de que $\alpha=\beta$, $g(\theta)=\alpha+\beta$, y si se contrasta la hipótesis de integrabilidad en la varianza según un IGARCH(1,1), en la que $\alpha+\beta=1$, entonces $g(\theta)=\alpha+\beta-1$.*

Por otro lado, y si queremos estudiar la discrepancia entre los parámetros definidos bajo H_0: $\theta=\theta_0$ (luego $\hat{\theta}-\theta_0$, es grande o pequeña) podemos utilizar el contraste de Wald (W), que basado en las estimaciones del modelo restringido[4], puede distinguirse entre robusto (W_R) y no robusto (W_R).

Asintóticamente, W, RV y LM son equivalentes. En muestras pequeñas su comportamiento es variable y su elección se debe a la facilidad de computación. En el caso del modelo de regresión lineal $LM \leq RV \leq W$, y no se producirá ninguna contradicción si por ejemplo LM rechaza la H_0 y W acepta, dada la mayor potencia de W.

4.1.2.- Modelos no anidados. Test de abarcamiento

La idea de este tipo de contrastes es propuesta en la literatura por Deaton (1982), Dastoor (1983), Mizon y Richard (1986)([5]). Formalmente, se trata de contrastar en términos lineales el modelo de volatilidad bajo la hipótesis nula con respecto a un

---

[4] *Habitualmente, el contraste de* H_0: $R\theta=r$ *para restricciones lineales es* $W = (R\hat{\theta}-r)'(R\Sigma^{-1}R')^{-1}(R\hat{\theta}-r) \sim \chi^2_q$ *donde* $var(R\hat{\theta}-r) = R\Sigma^{-1}R'$, *y* Σ *es la matriz de varianzas y covarianzas de los parámetros estimados y q el número de restricciones en la hipótesis nula. En el caso de restricciones lineales* $var(R\hat{\theta}-r) \neq R\Sigma^{-1}R'$, *por lo que*

$W = (R\hat{\theta}-r)'(var(R\hat{\theta}))^{-1}(R\hat{\theta}-r) \sim \chi^2_q$

siendo

$$var(R\hat{\theta}) = \left(\frac{\partial R\theta}{\partial \theta}\right) var(\hat{\theta}-\theta)\left(\frac{\partial R\theta}{\partial \theta}\right)'$$

donde $var(\hat{\theta}-\theta) = var(\hat{\theta}) = \sigma^2(X'X)^{-1}$, *similar a MCO. Estos contrastes se basan en que si* H_0 *es cierta, el vector discrepancia que se evalúa por máxima verosimilitud debería estar próximo a cero, ya que dicho estimador es consistente. El rechazo de* H_0 *implica que* $R\hat{\theta}_{MV}-r$ *es elevada.*

[5] *Se trata de contrastar dos modelos a partir de la regresión de una variable explicativa Y, y las variables explicativas de dos modelos: un modelo* M_1 *definido para las variables X, y otro modelo,* M_2, *definido para las variables Z, de tal forma que*

$$Y = X\beta + Z\delta + U$$

donde el modelo M_1 *será preferible al modelo* M_2, *cuando se cumpla la hipótesis en la que* H_0: $\delta=0$. *El estadístico aplicable es* $F(\delta=0) < F_{q,n-k}$, *donde q es el número de restricciones en* M_2 *y n-k los grados de libertad del modelo incluyendo todos los parámetros, k, de ambos modelos.*

conjunto de variables adicionales. Así,

$$h_t^2 = h_{0t}^2\left(\delta_0' y_{0t}\right) + \delta_a' y_{at} \qquad [4]$$

donde, $h_{0t}^2(\delta_0' y_{0t})$ es el modelo de volatilidad bajo la hipótesis nula (por ejemplo, un ARCH(1)); δ_0 es un vector kx1 de parámetros bajo H_0; y_{0t} es un vector kx1 de variables explicativas bajo H_0; δ_a es un vector mx1 de parámetros de las variables adicionales; y y_{at} es un vector mx1 de variables adicionales. Ya que el modelo de volatilidad bajo H_0 se obtiene cuando los parámetros de las variables adicionales, $_a y$, son ceros, se puede contrastar el modelo analizando si las restricciones de los parámetros adicionales son nulas, es decir, si $\delta_a=0$.

Por ejemplo, el contraste del modelo GARCH(1,1) frente al EGARCH(1,1) puede realizarse mediante el siguiente test de abarcamiento,

$$\log h_t^2 = \log\left(\omega + \beta h_{t-1}^2 + \alpha \varepsilon_{t-1}^2\right) + \beta^* \log h_{t-1}^2 + \gamma^* \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}^2}} + \alpha^* \left(\frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}^2}} - \sqrt{\frac{2}{\pi}} \right)$$

así, bajo la hipótesis nula, $\omega=\beta=\alpha=0$ el modelo es un EGARCH(1,1), mientras que si $\beta^*=\gamma^*=\alpha^*=0$ se acepta el modelo GARCH(1,1).

4.2.- Contrastes de estabilidad de las varianzas

La econometría tradicional tiene en los diferentes contrastes de Chow (entre otros su versión para la capacidad predictiva) un método de análisis de la estabilidad de los parámetros ante la posible existencia de un cambio de estructura. Recientemente, Chia-Shiang (1995) ha desarrollado una serie de contrastes de cambio estructural para las varianzas heterocedásticas, partiendo de un modelo GARCH(p,q) gaussiano. Se contrasta la igualdad de parámetros en dos submuestras. Así, la hipótesis planteada implica que

$$H_0: \theta_1 = \theta_2$$
$$H_a: \theta_1 \neq \theta_2$$

donde, $\theta_i=(\alpha_{i0}, \alpha_{i1},..., \alpha_{iq}, \beta_{i1}, \beta_{i2},..., \beta_{iq})'$ en el GARCH(p,q). θ_1 es el vector de parámetros de la primera submuestra y θ_2 es el vector de la segunda submuestra.

La justificación teórica del cambio estructural para la varianza condicionada puede encontrarse en la existencia de efectos integrados (por ejemplo, IARCH, IGARCH, entre otros).

4.2.1.- Contraste de estabilidad cuando δ es conocida

La aplicación del contraste necesita previamente del conocimiento del punto de ruptura, o momento donde se produce el cambio de estructura. Sea $[T\delta]$ la parte entera de $T\delta$, $0<\delta<1$, que es el punto de ruptura en el intervalo (0,1), entonces la función de verosimilitud puede particionarse según la siguiente expresión

$$L\left(y_1,...,y_T,h_t^2;\theta_1,\theta_2,\delta\right)=\frac{1}{T}\sum_{t-1}^{[T\delta]}\log f_{1t}+\frac{1}{T}\sum_{t-[T\delta]+1}^{T}\log f_{2t}$$

donde,

$$\log f_{it}=\frac{1}{2}\left(-\log h_{it}^2-\frac{\varepsilon_{it}^2}{h_{it}^2}\right),\quad h_{it}^2=\theta_i'Z_t$$

y

$$\theta_i=\left(\alpha_{i0},...\alpha_{iq},\beta_{i1},...\beta_{ip}\right),\quad i=1,2$$

$$Z_t=\left(1,\varepsilon_{t-1}^2,...\varepsilon_{t-q}^2,h_{t-1}^2,...h_{t-p}^2\right)$$

Teniendo en cuenta que r=(p+q+1) es el número de parámetros en la ecuación de la varianza, y $\theta=(\theta_1',\theta_2')'$, y basándonos en las puntuaciones (scores) de la función de

verosimilitud $L(\theta_1,\theta_2,\delta)$ cuando δ es conocida, podemos construir las condiciones de primer orden necesarias para la maximización de la función de verosimilitud.

Estas condiciones son iguales a

$$\frac{\partial L}{\partial \theta_1} = \frac{1}{T}\sum_{t=1}^{[T\delta]} \frac{1}{2} h_{1t}^{-2}\left(\frac{\partial h_{1t}^2}{\partial \theta_1}\right)\left(\frac{\varepsilon_t^2}{h_{1t}^2} - 1\right)$$

$$\frac{\partial L}{\partial \theta_2} = \frac{1}{T}\sum_{t=[T\delta]+1}^{T} \frac{1}{2} h_{2t}^{-2}\left(\frac{\partial h_{2t}^2}{\partial \theta_2}\right)\left(\frac{\varepsilon_t^2}{h_{2t}^2} - 1\right)$$

A partir de estas condiciones podemos realizar el contraste de la estabilidad de las varianzas basándonos en el multiplicador de Lagrange. Este, como hemos visto, requiere de la estimación de los parámetros bajo la hipótesis nula, H_0, cuando $\theta_1=\theta_2=\theta_0$. En este sentido, cuando δ es conocido el estadístico se construye según

$$LM(\delta) = T\left[\frac{\partial L(\delta)}{\partial \theta'}\right] B_0^{-1}\left[\frac{\partial L(\delta)}{\partial \theta'}\right] \simeq \chi^2_{p+q+1}$$

donde

$$B_0^{-1} = \frac{1}{4}\begin{bmatrix} \delta \nu_0 & 0 \\ 0 & (1-\delta)\nu_0 \end{bmatrix}$$

siendo v_0 definida por la siguiente expresión

$$v_0 = h_{0t}^{-2}\left(\frac{\partial h_{0t}^2}{\partial \theta_1}\right)\left(\frac{\varepsilon_t^2}{h_{0t}^2} - 1\right) \quad , \quad h_{0t}^2 = \theta_0' Z_t$$

y la decisión de aceptar la hipótesis nula sucede cuando LM $(\delta) > \chi^2_{p+q+1}$, donde p+q+1.

4.2.2.- Contraste de estabilidad cuando δ es desconocida

En este caso el estadístico se distribuye de forma diferente al caso [4.2.1] debido a que cuando δ es desconocido existe una convergencia débil en el proceso h^2_{0t} bajo la hipótesis nula. Así, el estadístico empleado será $LM(\delta) \sim B_r^2$, donde B_r^2 es la función de Bessell de orden r([6]) [véase Chia-Shiang (1995)].

De esta forma, los contrastes pueden reformularse de distintas maneras cuando δ es desconocida. En primer lugar, cuando la especificación de la varianza es correcta, los residuos estimados se distribuyen normalmente (por lo que se justifica que la función de verosimilitud sea normal), y además, $\{v_t\}$ es una martingala en diferencias (es decir, $E_{t-1}[v_t]=0$), permite replantear el multiplicador de Lagrange postulado en caso general, según

$$LMI(\delta) = T\left[\frac{\partial L(\delta)}{\partial \theta'}\right]\left(A_0^{-1}(\delta)\right)\left[\frac{\partial L(\delta)}{\partial \theta'}\right] \sim B_r^2$$

[6] *Este hecho se deriva del teorema de posicionamiento continuo de Billingsley (1968) donde* $\sup_{\delta \in \pi} LM(\delta) \Rightarrow \sup_{\delta \in \pi} B_r^2(\delta)$, y $\pi \subset [0,1]$.

donde una estimación consistente de A_0 es

$$\hat{A}_T(k) = \frac{1}{2}\begin{bmatrix} \frac{k}{T}\hat{G}_T & 0 \\ 0 & \left(1-\frac{k}{T}\right)\hat{G}_T \end{bmatrix}$$

en la que, k/T es un valor que discretiza δ, siendo k un número entero que se identifica con el punto de corte. Por otro lado, el estimador consistente de G_0 está definido por

$$\hat{G}_T = \frac{1}{T}\sum_{t=1}^{T} h_t^4(\hat{\theta})\left[\frac{\partial h_t^2(\hat{\theta})}{\partial \theta}\right]\left[\frac{\partial h_t^2(\hat{\theta})}{\partial \theta}\right]'$$

Los valores críticos de la distribución B_r^2 están tabulados en la Tabla I (pág. 249) para distintos valores de r y probabilidad $P[B_r^2 > c_r(\alpha)]=\alpha$ [véase Chia-Shiang (1995)][7].

Sin embargo, Chia-Shiang (1995) especifica diferentes contrastes sobre la varianza que se basan en distintos escenarios. Por un lado, considera que la varianza condicional está correctamente especificada y que los residuos son no normales (estadístico LM2). Por otro lado, que la especificación de la varianza es incorrecta (estadístico LM3), finalmente, el contraste de la varianza no condicionada (estadístico LMU) (véase el cuadro 10). En comparación con el estadístico LM1, éste parece tener la misma distribución límite que el estadístico LM2.

[7] *Algunos puntos críticos para diversos niveles de significación son:*

$$c(0.1) = 7.728 \ , \ c(0.05) = 9.303 \ , \ c(0.01) = 12.888$$

Por lo que se refiere al estadístico LM3, éste se utiliza cuando el modelo para la varianza heterocedástica está incorrectamente especificado. En este caso, resulta inválido el empleo del estadístico LM2, porque la estimación consistente de v_0 (v_T, es el estimador de la matriz de covarianzas de los estimadores) no es la correcta. De esta forma, la estimación de v_0 se realizará utilizando un procedimiento consistente ante la existencia de heterocedasticidad y autocorrelación [método de Newey y West (1987)] y empleando el kernel de Barttlet (si bien pueden utilizarse otros kernels como Parzen, etc.).

Cuadro 9.- Procedimiento de constraste cuando δ es desconocida.

a) Estimar el vector de parámetros θ por BHHH, generando las series $\left\{\frac{\partial h_t^2(\hat{\theta})}{\partial \theta}\right\}$ y $\left\{h_t^2(\hat{\theta})\right\}$, con t=1,...,T.

b) Calcular la estimación de v_T.

c) Determinar el punto de corte $k=k^*, k^*+1,..., n-k^*$, siendo $k^*=[0.1T]$ la parte entera de 0.1T. Para cada k se obtienen las sumas parciales estimadas S_1 y S_{T-k}.

d) Para cada k, se calcula el estadístico LM.

e) Si el máximo de LM(k/T) es mayor que el valor crítico $c_T(\alpha)$, se rechaza la hipótesis nula de igualdad de varianzas, y por tanto de acepta el cambio estructural al nivel de significación estadística α, sabiendo que los grados de libertad son r=p+q+1

Sin embargo, previamente debe ajustarse un modelo autorregresivo de orden uno (AR(1)) para cada elemento de

$$\hat{v}_t = h_t^{-2}(\hat{\theta})\left(\frac{\partial h_t^2(\hat{\theta})}{\partial \theta_1}\right)\left(\frac{\varepsilon_t^2}{h_t^2(\hat{\theta})} - 1\right)$$

A partir de ellos, obtenemos el parámetro ρ_i y *la varianza del error es* σ_i^2, $i=1,...,r$

Cuadro 10.- Distintos estadísticos de cambio estructural.

Estadístico	v_T	S_k	S_{T-k}	LM
LM2	$\hat{v}_T = \frac{1}{T}\sum_{t=1}^{T} \hat{h}_t^4 \left[\frac{\partial h_t^2(\hat{\gamma})}{\partial \gamma}\right]\left[\frac{\partial h_t^2(\hat{\gamma})}{\partial \gamma}\right]' \left(\frac{\varepsilon_t^2}{\hat{h}_t^2} - 1\right)$	$\hat{S}_k = \sum_{t=1}^{k} h_t^{-2}(\hat{\theta})\left(\frac{\partial h_t^2(\hat{\theta})}{\partial \theta}\right)\left(\frac{\varepsilon_t^2}{h_t^2(\hat{\theta})} - 1\right)$	$\hat{S}_{T-k} = \sum_{t=k+1}^{T} h_t^{-2}(\hat{\theta})\left(\frac{\partial h_t^2(\hat{\theta})}{\partial \theta}\right)\left(\frac{\varepsilon_t^2}{h_t^2(\hat{\theta})} - 1\right)$	$LM2(k/T) = \frac{1}{k}\hat{S}_k' \hat{v}_k^{-1} \hat{S}_k + \frac{1}{T-k}\hat{S}_{T-k}' v_T^{-1} \hat{S}_{T-k}$
LM3	$\hat{v}_T = \frac{T}{T-r}\sum_{j=-T+1}^{T-1} K\left(\frac{j}{M^*}\right)\Gamma(j)$, $K(x) = 1 - \lvert x \rvert$ $\Gamma(j) = \frac{1}{T}\sum_{t=j+1}^{T} \hat{v}_t \hat{v}_{t-j}$, $j \geq 0$ $= \frac{1}{T}\sum_{t=-j+1}^{T} \hat{v}_{t+j} \hat{v}_t$, $j < 0$	$\hat{S}_k = \sum_{t=1}^{k} h_t^{-2}(\hat{\theta})\left(\frac{\partial h_t^2(\hat{\theta})}{\partial \theta}\right)\left(\frac{\varepsilon_t^2}{h_t^2(\hat{\theta})} - 1\right)$	$\hat{S}_{T-k} = \sum_{t=k+1}^{T} h_t^{-2}(\hat{\theta})\left(\frac{\partial h_t^2(\hat{\theta})}{\partial \theta}\right)\left(\frac{\varepsilon_t^2}{h_t^2(\hat{\theta})} - 1\right)$	$LM3(k/T) = \frac{1}{k}\hat{S}_k' \hat{v}_k^{-1} \hat{S}_k + \frac{1}{T-k}\hat{S}_{T-k}' v_T^{-1} \hat{S}_{T-k}$
LMU	$\{\hat{v}_t\} = \{\hat{\varepsilon}_t - \hat{\sigma}_u^2\}$, $t = 1,\ldots,k$ $\hat{\sigma}_u = \sum_{t=1}^{T} \frac{\varepsilon_t^2}{T}$	$\hat{S}_k = \sum_{t=1}^{k} \hat{v}_t$	$\hat{S}_{T-k} = \sum_{t=k+1}^{T} \hat{v}_t$, $k = [0.1T], \ldots, [0.9T]$	$LMU(k/T) = \hat{\sigma}_v^{-2}\left(\frac{1}{k}\hat{S}_k^2 + \frac{1}{T-k}\hat{S}_{T-k}^2\right)$

Luego, se calcula una estimación de ϕ

$$\hat{\phi} = \frac{\sum_{i=1}^{r} \frac{1}{r} \frac{4\hat{\rho}_i^2 \hat{\sigma}_i^4}{(1-\hat{\rho}_i)^6 (1+\hat{\rho}_i)^2}}{\sum_{i=1}^{r} \frac{1}{r} \frac{\hat{\sigma}_i^4}{(1-\hat{\rho}_i)^4}}$$

y se determina el truncamiento según el valor de M* definido por $1.1447(\phi T)^{1/3}$.

Por último, podemos referirnos a la expresión utilizada en el contraste de estabilidad de la varianza no condicionada (LMU). La expresión LMU(k/T) se obtiene una vez se ha calculado el estimador consistente de la varianza, según la transformación de Newey y West (1987), tal que

$$\hat{\sigma}_v^2 = \frac{1}{T}\sum_{t=1}^{T} \hat{v}_t^2 + 2\sum_{\tau=1}^{g} \omega_{\tau g} \sum_{t=\tau+1}^{T} \hat{v}_t \hat{v}_{t-j} \ , \quad \omega_{\tau g} = 1 - \left[\frac{\tau}{g+1}\right]$$

donde g es el parámetro de truncamiento que puede determinarse a través del procedimiento de ventana automática de Andrew (1991). Si el máximo de LMU(k/T) es mayor que $c(\alpha)$, entonces se rechaza la hipótesis nula, $H_0: \sigma_u^2 = \sigma^2$, siendo σ^2 la varianza constante.

A diferencia del resto de estadísticos, este test se basa en Pagan y Schwert (1990), y es más simple que los anteriores, porque sólo requiere de la estimación por mínimos cuadrados ordinarios y la estimación consistente de Newey y West para σ_v^2. Una de las condiciones para que el contraste LMU sea correcto es que los momentos de cuarto orden sean finitos para los datos normalizados o estandarizados, $\epsilon_t^2/\sqrt{h^2_{0t}}$. Si no fuese así, el contraste sería inconsistente.

BIBLIOGRAFIA

- BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity". Journal of Econometrics, Vol. 31, págs. 307-327.

- BROCK, W.A., DECHER, W.D. and SHEINKMAN, J.A. (1987): "A Test for Independence based on correlation dimension". Working Paper. University of Wisconsin.

- CHIA-SHIANG, J.C. (1995): "Detecting Parameter Shift in Garch models". Econometric Reviews, Vol. 14, págs. 241-266.

- DASTOOR, N.K. (1983): "Some Aspects of Testing Non-Nested Hypotheses". Journal of Econometrics. Vol. 21, págs. 213-228.

- DEATON, A.S. (1983): "Model Selection Procedures, or, Does the Consumption Function Exist?". In Evaluating the Realibility of Macroeconomic Models, eds. G.C.Chow and P.Corsi. New York. John Whiley.

- ENGLE. R.F. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of Variance of U.K. Inflation". Econometrica, Vol. 50, págs. 987-1007.

- ENGLE, R.F., CHOU, R. AND KRONER, K. (1992): "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence". Journal of Econometrics, Vol. 50, págs. 5-59.

- ENGLE, R.F. AND NG, V. (1993): "Measuring and Testing the impact of News on Volatility". Journal of Finance, Vol. 48, págs. 1749-1778.

- HAMILTON, J.D. AND SUSMEL, R. (1994): "Autoregressive Conditional Heterokedasticity and Changes in Regime". Journal of Econometrics. Vol. 64, págs. 307-333.

- HOLT, M. AND ARA1DHYULA, S. (1990): "Price Risk in Supply Equations: An Application of GARCH Time Series Models to the U.S. Broiler Market". Southern Economic Journal. Vol. 57, págs. 230-240.

- HSU, D. (1982): "A Bayesian Robust Detection of Shift in the Risk Structure of Stock Market Returns". Journal of the American Statistical Association, Vol. 77, págs.
29-39.

- LUMSDAINE, R.L. (1993): "Asymptotic Properties of the Quasi-Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models". Unpublished Manuscript, Priceton University, Department of Economics.

- LUMSDAINE, R.L. (1995): "Finite-Sample Properties of the Maximum Likelihood Estimator in Garch(1,1) and Igarch(1,1) models: A Monte Carlos Investigation". Journal of Business and Economics Statistics, *Vol. 13, págs. 1-10.*

- MIZON, G.F., AND RICHARD, J.F. (1986): "The Encompassing Principle and its Application to Non-Nested Hypotheses". Econometrica. Vol. 54, págs. 657-678.

- NEWEY, W.K. AND WEST, K.D. (1987): "A Simple, Positive, Semi-Definite Heterokedasticity and Autocorrelation Consistent Covariance Matrix". Econometrica, Vol. 55, págs. 703-708.

- PAGAN, A.R. AND SCHWERT, G.W. (1990): "Alternative Models for Conditional Stock Volatility". Journal of Econometrics, Vol. 45, págs.267-290.

- PÉREZ-RODRÍGUEZ, J.V. (1995a): "El Modelo ARCH Uniecuacional". Documento de Trabajo, DOCT-95D01. Departamento de Econometría, Estadística y Economía Española.

- PÉREZ-RODRÍGUEZ, J. (1995b): "Extensiones del Modelo ARCH: Modelos ARCH Uniecuacionales y Multiecuacionales". Documento de Trabajo, DOCT-95D02. Departamento de Econometría, Estadística y Economía Española.

- WEISS, A.A. (1986): "Asymptotic Theory for ARCH Models: Estimation and Testing". Econometric Theory, Vol. 2, págs. 107-131.