

Análisis de la volatilidad de series financieras mediante modelos ARMA-GARCH

Trabajo Final de Grado

Autor: Sandra Ferrando Latorre

Director: Manuel Salvador Figueras

Codirector: Jesús Miguel Álvarez

Índice

Resumen.....	2
1. Introducción.....	3
2. Modelos ARMA-GARCH.....	5
2.1. Planteamiento del problema.....	5
2.2. Extensiones del modelo: modelos APARCH.....	8
3. Estimación y selección de los modelos ARMA-GARCH.....	9
3.1. Estimación del modelo.....	9
3.2. Identificación del modelo.....	10
4. Aplicación empírica: análisis de índices bursátiles.....	11
4.1. Análisis estadístico del IBEX35.....	11
4.2. Resultados del proceso de estimación y selección de modelos para el resto de índices.....	23
5. Validación predictiva extramuestral de los modelos estimados.....	26
5.1. Predicción en modelos ARMA-GARCH.....	28
5.2. Criterios de comparación.....	28
5.2.1. Evaluación del comportamiento predictivo puntual.....	28
5.2.2. Evaluación del comportamiento predictivo por intervalos.....	30
5.3. Resultados obtenidos.....	31
5.3.1. IBEX35.....	31
5.3.2. Resto de los índices.....	34
6. Conclusiones.....	35
7. Bibliografía.....	36

Resumen

En este trabajo se estima la volatilidad de un conjunto de índices bursátiles analizando su impacto sobre la predicción extramuestral. Para ello se utilizan modelos ARMA-GARCH y APARCH, analizando el impacto ejercido por la volatilidad, la existencia de efecto asimétrico y la falta de normalidad de la distribución del error. En el trabajo se describe cómo identificar, estimar y analizar la bondad de ajuste de este tipo de modelos así como elaborar predicciones extramuestrales. Así mismo se realiza una validación predictiva extramuestral de este tipo de modelos, tanto a nivel puntual como a nivel de intervalos, utilizando el método rolling. La metodología se aplica a 6 índices bursátiles. Los resultados obtenidos muestran que los modelos heteroscedásticos tienen un mejor rendimiento a nivel de intervalos, mostrando un comportamiento más adaptativo a las oscilaciones de la serie, sin que aprecien diferencias significativas en las predicciones elaboradas a nivel puntual debido la eficiencia de los mercados. Por su parte la incorporación del efecto asimétrico tiene un carácter más marginal y no apreciándose mejora sistemática alguna en la utilización de errores no normales.

Palabras clave: GARCH, APARCH, Máxima verosimilitud, Validación predictiva extramuestral, Método Rolling, Series financieras, Índices bursátiles

Abstract

In this work the volatility of a set of stock indices is estimated by means of ARMA GARCH and APARCH models and its impact on the out-sampling prediction is analyzed, as well as the existence of asymmetric effect and the lack of normality of the distribution of the error. The work describes how to identify, estimate and analyze the goodness of fit of these models, and to build one-step forward predictions. Also an out-sampling predictive validation is carried out in terms of point and intervals predictions and using a rolling method. The methodology is applied to 6 stock indices. The results show that point predictions are very similar for all the models due to the efficiency of markets. With respect to prediction by intervals the results show that the heteroscedastic models have better performance, by showing a more adaptive behavior to oscillations in the return series. The incorporation of an asymmetric effect in the model, even though statistically significant, has a marginal impact and it is not appreciated significant differences in the use of non-normal distributions.

Keywords: GARCH, APARCH, Maximum likelihood, Outsampling predictive validation, Rolling Method, Financial series, Stock indices.

1. Introducción

En los últimos años los estudios acerca de la volatilidad y su modelización, tanto empírica como teórica, han ido ganando peso en el ámbito de los mercados económicos y financieros debido al papel fundamental que juega en la evaluación del riesgo asociado a los activos financieros (ver, por ejemplo, Tsay, 2013 capítulos 4 y 5). La volatilidad de un activo financiero podría definirse como la intensidad o frecuencia de los cambios que se producen en sus precios y suele medirse mediante medidas de variabilidad de su rentabilidad.

En la práctica podemos hablar de la existencia de varios tipos de volatilidad como son la volatilidad pasada o histórica, la volatilidad implícita, la volatilidad observada o la volatilidad condicional o futura. La volatilidad histórica se puede definir como una medida retrospectiva de la volatilidad, la cual refleja el comportamiento de la variabilidad de los precios en el pasado. La volatilidad implícita está ligada a la valoración de opciones financieras que son instrumentos que ofrecen a sus propietarios el derecho a comprar o vender un activo determinado a un precio fijo en algún momento en el futuro. En el mercado se utiliza de forma general la fórmula de Black-Scholes que permite expresar la volatilidad implícita utilizando como dato de entrada la prima de la opción que se está negociando en el mercado. La volatilidad observada se calcula a partir de la varianza de la rentabilidad del activo observada con alta frecuencia (cada 5 o 10 minutos). Finalmente la volatilidad condicional o futura se define como la varianza o desviación típica de la rentabilidad futura del activo la cual se calcula a partir de modelos econométricos siendo los modelos ARMA-GARCH y los modelos de volatilidad estocástica los más utilizados.

Ambos tipos de modelos estiman la volatilidad mediante la varianza o la desviación típica condicional de la serie de rentabilidades, la cual los modelos ARMA-GARCH modelan como función exacta de las innovaciones del modelo y sus retardos mientras que los modelos de volatilidad estocástica utilizan procesos estocásticos.

La familia de modelos GARCH fue introducida por Engle (1982) que definió los modelos ARCH (iniciales en inglés de Condicionalmente Heteroscedástico y Autorregresivo) cuya varianza condicional es función lineal de los cuadrados de las innovaciones del modelo así como sus retardos dando lugar a un proceso autorregresivo (AR) en los cuadrados de dichas innovaciones. Los modelos ajustados resultaron ser

poco parsimoniosos precisando inclusión de un número elevado de retardos debido a la existencia del fenómeno “*volatility clustering*” en el que la volatilidad se agrupa en bloques de alta persistencia temporal. Con el fin de resolver dicho problema Bollerslev (1986) propone los modelos GARCH (iniciales en inglés de Condicionalmente Heteroscedástico Autorregresivo Generalizado) que incluye, el ecuación de evolución de la varianza condicional, valores retardados de la misma lo cual permite modelar de forma más parsimoniosa su evolución, llegando a ser el modelo GARCH (1,1) un modelo estándar en la literatura.

Como Tsay (2013) argumenta, los modelos ARMA-GARCH presentan una serie de debilidades que debemos tener en cuenta:

- Suponen que los shocks positivos y negativos tienen el mismo tipo de efecto sobre la volatilidad por lo que son incapaces de capturar el llamado efecto asimétrico.
- Imponen bastantes restricciones sobre los parámetros para tener momentos finitos.
- No proporcionan información acerca de las fuentes de variación de las series financieras.
- Tienden a sobre-predecir la volatilidad debido a que responden lentamente a shocks grandes que se producen de forma aislada.
- Las estimaciones obtenidas suelen tener alta persistencia y leptocurtosis condicional y, en algunos casos, asimetría en la distribución de los errores

Para tratar dichos problemas se han propuesto diversas extensiones en la literatura: los modelos EGARCH de Nelson (1991), TGARCH de Glosten y otros (1993), NGARCH de Engle y Ng (1993), APARCH de Ding y otros (1993) o el enfoque semiparamétrico de Engle y González-Rivera (1991) entre muchos otros.

Por su parte, los modelos de volatilidad estocástica modelan la varianza condicional de la serie de rentabilidades utilizando procesos estocásticos siendo el proceso AR (1) de Taylor (1994) el más utilizado. Este tipo de modelos es más difícil de estimar debido al carácter latente de la volatilidad, siendo su ajuste muy similar al de los modelos de la familia GARCH. Por dicha razón nos centraremos en esta última familia.

En este trabajo se realiza una estimación de la volatilidad de 6 índices bursátiles internacionales entre los que se encuentra el IBEX35, y se evalúa su impacto en la predicción de su rentabilidad. Se analiza, además, la existencia de efecto asimétrico y la falta de normalidad de las innovaciones del modelo utilizando diversos criterios propuestos en la literatura para evaluar dichas predicciones en términos puntuales y por intervalos.

En la sección 2 se realiza una descripción de los elementos principales en los modelos ARMA-GARCH, describiendo además el efecto llamado persistencia en volatilidad captada por el modelo GARCH, así como de su variante APARCH utilizada para capturar y analizar la existencia de efecto asimétrico. La sección finaliza con una descripción de los criterios utilizados para evaluar las predicciones del modelo. En la sección 3 se explica cómo ajustar un modelo ARMA-GARCH a un conjunto de datos: cómo identificarlo, estimarlo, analizar su bondad de ajuste y utilizarlo para predecir. En la sección 4 se aplicar la metodología descrita en la sección anterior al análisis de 6 índices bursátiles prestando especial atención al IBEX35. En la sección 5 se realiza un estudio comparativo en términos predictivos, de los modelos ajustados en la sección 4 y, finalmente, la sección 6 concluye destacando los hechos más relevantes encontrados en el trabajo. Se incluyen además 6 anexos en los que se muestran con más detalles los resultados obtenidos en el proceso de estimación y selección de modelos llevado a cabo para cada uno de los índices anteriores.

2. Modelos ARMA-GARCH

2.1. Planteamiento del problema

Sea $\{P_t; t = 1, \dots, T\}$ la serie de precios de un activo financiero donde P_t es el precio del activo en el periodo t .

La mayoría de los estudios financieros involucran rentabilidades en lugar de precios. Esto es así porque para la mayoría de inversores el rendimiento de un activo es un resumen completo y libre de escala de la oportunidad de inversión; además, las series de rentabilidades son más fáciles de manejar que las series de precios, puesto que las primeras tienen propiedades más atractivas como, por ejemplo, la estacionariedad, que facilitan su tratamiento estadístico.

La rentabilidad de un activo financiero se puede calcular de dos formas: simple o continua. La rentabilidad simple es la tasa de crecimiento relativa del nivel de precios

del activo en un periodo y viene dada por $R_t = 100 \frac{P_t - P_{t-1}}{P_{t-1}}$; por su parte la rentabilidad

continua se define como la tasa de crecimiento acumulativa media en un periodo cuando éste se divide en intervalos de tiempo infinitesimales y viene dada por la expresión

$r_t = 100 \log \left(\frac{P_t}{P_{t-1}} \right)$ verificándose que $P_t = \lim_{n \rightarrow \infty} P_{t-1} \left(1 + \frac{r_t}{100n} \right)^n$. En este trabajo

utilizaremos ésta última dado que su tratamiento estadístico es más ventajoso. Esta ventaja se derivan de que la rentabilidad continua para un horizonte $h > 0$ verifica que

$r_t(h) = 100 \log \left(\frac{P_{t+h}}{P_{t-1}} \right) = r_t + \dots + r_{t+h-1}$ lo cual facilita la elaboración de predicciones a

varios horizontes.

Nuestro objetivo es construir un modelo que describa la evolución de las rentabilidades $\{r_t; t = 2, \dots, T\}$ a lo largo del tiempo así como de su volatilidad $\sigma_t^2 = \text{Var}(r_t | \Omega_{t-1})$ donde $\Omega_t = \{r_1, \dots, r_t\}$ es la información disponible al analista en el periodo t . Para ello utilizaremos un ARMA(r,s)-GARCH(p,q) viene dado por las expresiones:

$$\phi(B)(r_t - \mu) = \theta(B)\varepsilon_t$$

$$\beta(B)\sigma_t^2 = \omega + \alpha(B)\varepsilon_t^2$$

con $\varepsilon_t = \sigma_t z_t$ y $z_t | \Omega_{t-1} \sim D(0, 1, \lambda)$ donde $D(0, \sigma_t, \lambda)$ denota una distribución con media 0, desviación típica 1 y vector de parámetros λ .

B es el operador retardo

$\mu = E[r_t]$ media incondicional de la serie r_t

$\phi(z), \theta(z), \alpha(z)$ y $\beta(z)$ polinomios de grados r, s, p y q , respectivamente

$\phi(z), \theta(z)$ sin raíces comunes todas ellas fuera del círculo unidad

$\omega > 0$

$\alpha(z)$ y $\beta(z)$ sin raíces comunes y coeficientes no negativos

$\alpha(1) + \beta(1) < 1$ con el fin de garantizar que la varianza incondicional del modelo existe de forma que $\sigma^2 = \text{Var}(r_t) = \frac{\omega}{1 - \alpha(1) - \beta(1)}$.

En este caso $\alpha(1) + \beta(1)$ mide la persistencia en volatilidad del proceso de forma que cuanto mayor es su valor mayor, mayor es la duración del impacto que ejerce una innovación ε_t en la volatilidad del proceso y, por tanto, mayor es el fenómeno de “*volatility clustering*” de forma que, si en un periodo la volatilidad ha sido alta en el periodo siguiente también tenderá a ser alta, y si por el contrario, la volatilidad es pequeña en un periodo, también será pequeña en el periodo siguiente.

Respecto a la distribución $D(0,1, \lambda)$ en la literatura se han propuesto diversas formas de la misma siendo las más utilizadas la distribución normal estándar $N(0,1)$, la distribución t de Student centrada y tipificada $\sqrt{\frac{v}{v-2}} t_v$ con $v > 2$ grados de libertad o la distribución exponencial generalizada (GED) $GED(0,1,v)$ con v grados de libertad propuesta por Nelson (1991). Sus funciones de densidad vienen dadas por las expresiones:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \text{ en el caso de la distribución } N(0,1)$$

$$f(x) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{(v-2)\pi}} \left(1 + \frac{x^2}{v-2}\right)^{-\frac{v+1}{2}} \text{ en el caso de la } \sqrt{\frac{v}{v-2}} t_v$$

$$f(x) = \frac{v e^{-\frac{1}{2}|x|^v}}{2^{\frac{1+1}{v}} \Gamma\left(\frac{1}{v}\right)} \text{ en el caso de la } GED(0,1,v)$$

Las 3 distribuciones son simétricas en torno a 0. Además, la distribución t de Student es leptocúrtica y su leptocurtosis es más elevada cuanto menor es el valor de v . Algo similar ocurre con la distribución GED que es leptocúrtica si $v < 2$, normal si $v = 2$ y platicúrtica si $v > 2$.

En algunas series se observa, además, la presencia de una asimetría (habitualmente negativa) no muy pronunciada en la distribución de las innovaciones.

Para capturar y analizar la significación de dicho efecto utilizaremos las versiones asimétricas de las 3 distribuciones anteriores, $SN(0,1,\xi)$, $St_v(0,1,\xi)$ y $SGED(0,1,v,\xi)$ cuya función de densidad viene dada por la expresión genérica:

$$g(x) = \begin{cases} \frac{2}{\xi + \frac{1}{\xi}} f(x) & \text{si } x < 0 \\ \frac{2}{\xi + \frac{1}{\xi}} f\left(\frac{x}{\xi}\right) & \text{si } x \geq 0 \end{cases}$$

siendo f la función de densidad de la distribución original. Estas distribuciones son simétricas en torno al 0 si $\xi=1$, asimétricas a izquierda si $\xi<1$ y a derechas si $\xi>1$.

2.2. Extensiones del modelo: modelos APARCH

Los modelos GARCH suponen que el impacto ejercido por las innovaciones ε_t es independiente de su signo. El trabajo en el modelado de la volatilidad condicional de la serie de precios de las acciones ha demostrado, en general, que la volatilidad de un activo responde de forma diferente a los shocks positivos frente a los negativos de la misma cuantía. Por lo general, un shock negativo produce un efecto mayor en la volatilidad, que el generado por un shock positivo debido a que una sobrevaloración del precio de un activo tiende a incrementar la incertidumbre sobre su valor. Este efecto se conoce en la literatura con el nombre de “efecto asimétrico”.

Para tratar de capturar este tipo de efecto y analizar su significación estadística se han propuesto en la literatura diversas extensiones del modelo GARCH: modelo EGARCH (modelo GARCH Exponencial) propuesto por Nelson (1991); TGARCH (modelo GARCH Umbral, también conocido como modelo GJR) propuesto por Glosten y otros (1993) , NGARCH (modelo GARCH No Simétrico) propuesto por Engle y Ng (1993), APARCH (modelo condicional heteroscedástico autorregresivo de potencia asimétrica) propuesto por Ding y otros (1993), En este trabajo utilizaremos el modelo ARMA(r,s)-APARCH(p,q) cuya expresión viene dada por:

$$\phi(B)(r_t - \mu) = \theta(B)\varepsilon_t \text{ con } \varepsilon_t = \sigma_t z_t \text{ y } z_t | \Omega_{t-1} \sim D(0,1,\lambda)$$

$$\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta$$

La existencia de efecto asimétrico de los residuos sobre la volatilidad vendrá marcada por los parámetros γ de forma que si alguno de ellos es significativo marcará la existencia de un impacto diferente de los residuos negativos con respecto a los residuos positivos. En particular si $p=q=1$ el efecto asimétrico descrito habitualmente en la literatura existirá siempre y cuando $\gamma_1 > 0$.

3. Estimación y selección de modelos ARMA-GARCH

3.1. Estimación del modelo

La estimación de los modelos anteriores se realiza por el método de la máxima verosimilitud. Todos los modelos son casos particulares del siguiente

$$r_t = G(x_t; \zeta) + \sigma_t(x_t; \psi) z_t$$

$$z_t | \Omega_{t-1} \sim D(0, 1, \lambda)$$

donde

$$x_t = (1, r_{t-1}, \dots, r_{t-p}, \varepsilon_{t-1}, \dots, \varepsilon_{t-q})$$

$G(x_t; \zeta)$ es función de ζ dos veces continuamente diferenciable

$\sigma_t^2(x_t; \psi)$ es función de ψ dos veces continuamente diferenciable

El logaritmo de la función de verosimilitud viene dado por:

$$\ell_{\max\{r,s,p,q\}+1:T}(\zeta, \psi, \lambda) = \sum_{t=\max\{r,s,p,q\}+1}^T \ell_t(\zeta, \psi, \lambda)$$

donde $\ell_t(\zeta, \psi) = \log \left(\frac{1}{\sigma_t(x_t; \psi)} f \left(\frac{r_t - G(x_t; \zeta)}{\sigma_t(x_t; \psi)}; \lambda \right) \right)$ siendo f la función de densidad de la

distribución $D(0, 1, \lambda)$. El estimador máximo verosímil de (ζ, ψ, λ) se obtiene maximizando la función anterior y se puede probar que, bajo ciertas condiciones de regularidad la distribución del estimador máximo-verosímil es asintóticamente normal con media el verdadero valor de los parámetros anteriores $(\zeta_0, \psi_0, \lambda_0)$ y matriz de varianzas y covarianzas dada por $\frac{1}{T} \Sigma$, cuyo valor depende del gradiente y el hessiano

de la función de verosimilitud evaluado en el estimador máximo-verosimil. La matriz anterior es diagonal por bloques respecto a los 3 parámetros ζ , ψ y λ lo cual facilita la labor de identificación del modelo dado que la estimación de cada uno de los 3 parámetros es asintóticamente independiente del resto.

3.2. Identificación del modelo

Dado el resultado de la sección anterior la identificación del modelo se realiza en 3 pasos:

- a) Se identifican primero los órdenes r y s de la parte ARMA utilizando las herramientas clásicas de identificación este tipo de modelos: correlogramas, criterios de selección de modelos
- b) Una vez identificados r y s se estima el modelo ARMA(r,s) homoscedástico con errores normales por máxima verosimilitud y se calculan sus residuos cuadráticos $\hat{\varepsilon}_t^2$. Se analiza el correlograma de estos residuos y, a partir de ellos se identifica los órdenes p y q de un modelo GARCH. Habitualmente si el correlograma tiene una estructura de un AR(p) se plantea un modelo ARCH(p); en otro caso se plantea un modelo GARCH(1,1) que es el más simple
- c) Se estima el modelo ARMA(r,s)-GARCH(p,q) con errores normales y se calculan los residuos tipificados $\hat{z}_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_t}$. Si el modelo está bien especificado en media y varianza esta serie será ruido blanco débil homoscedástico. Si no lo es, se re-especifica el modelo ARMA-GARCH en cuanto a sus órdenes y se vuelve a repetir el paso c)
- d) Si $\hat{z}_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_t}$ es ruido blanco se analiza la hipótesis de normalidad y, en caso de ser rechazada se identifica cuál de las distribuciones descritas en la sección 2.1 se ajusta a los datos utilizando QQ plots o test de bondad de ajuste como el de Kolmogorov-Smirnov. Si hay varias que se ajustan bien se estiman los modelos ARMA(r,s)-GARCH(p,q) correspondientes a cada una de ellas y se selecciona aquél con mejor valor con respecto a algún criterio

de selección de modelos. En nuestro caso utilizaremos el criterio BIC que es consistente.

4. Aplicación empírica: análisis de índices bursátiles

En esta sección se estiman y seleccionan los modelos ARMA-GARCH que mejor se ajustan a los 6 índices bursátiles analizados en el trabajo aplicando la metodología descrita en la sección anterior. Los datos corresponden a las series de precios de cierres diarios de los índices BEL20, DAX, DOWJONES, EUROSTOXX, IBEX35 y NIKKEI desde el 02/01/2001 hasta el 24/04/2015 dando un total de 3644 observaciones por serie. En la sección se muestra con detalle el proceso seguido para el IBEX35. Los resultados para el resto de las series se muestran en los anexos correspondientes. Todos los resultados que se exponen en el trabajo fueron obtenidos utilizando librerías del programa R 3.1.2.

4.1. Análisis estadístico del IBEX35

Comenzamos nuestro estudio analizando las raíces unitarias del logaritmo de la serie de precios.

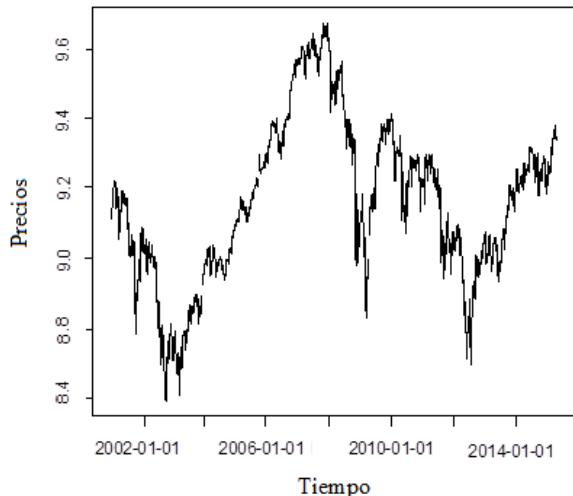


Figura 4-1: Gráfico de log precios contra el tiempo

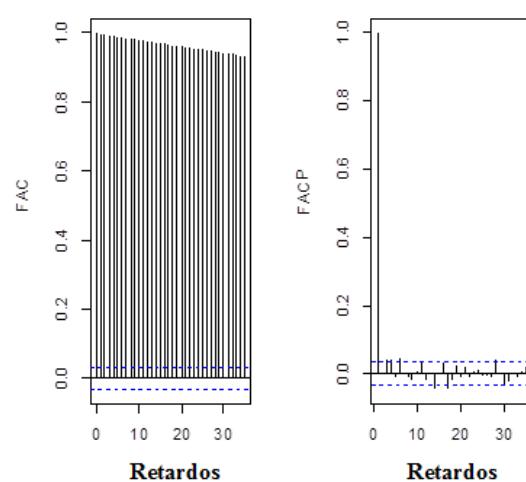


Figura 4-2: Correlograma log precios

A simple vista podemos observar en el gráfico que la serie del logaritmo de precios (Figura 4-1) presenta no estacionariedad, incluso podríamos afirmar que presenta tendencia estocástica y determinista, pues tenemos subperíodos grandes con pauta creciente o decreciente y, además, la dispersión de los valores es creciente ya que a medida que aumenta el tiempo la diferencia en precios es mayor.

Atendiendo al correlograma (Figura 4-2), vemos que el correspondiente a la función de autocorrelación (el de la izquierda, FAC) presenta un decrecimiento lento con valores próximos a 1 y, el correlograma de la función parcial de autocorrelación (el de la derecha, FACP) presenta una primera barra cercana a uno y el resto de valores oscilan en torno a cero. Volvemos a predecir que estamos ante una serie no estacionaria.

Para confirmar la no estacionariedad de la serie realizamos los dos contrastes principales: Dickey-Fuller y KPSS, ambos se estimaron con constante pero sin tendencia, pues ésta última resultó no significativa.

Para el contraste Dickey-Fuller consideramos la hipótesis nula H_0 : Existencia de raíces unitarias o No estacionariedad de la serie, frente la hipótesis alternativa H_a : No existencia de raíces unitarias o Estacionariedad de la serie. Mientras que para el contraste KPSS ocurre lo contrario: bajo la hipótesis nula el proceso es estacionario y, bajo la hipótesis alternativa presenta una raíz unitaria.

El resultado obtenido para el contraste Dickey-Fuller fue de un p.valor de 0.4427, atendiendo a dicho valor, para los niveles de significatividad del 1% y 5% resultaron que no se rechaza la hipótesis nula de existencia de raíces unitarias, siendo, por tanto, la serie no estacionaria. Así mismo, para el contraste KPSS el p.valor es menor que 0.01 y, por tanto, existe evidencia en contra de la hipótesis de estacionariedad para un nivel de significatividad del 1%. Por tanto, ambos tests proporcionan evidencia de que la serie de precios del IBEX35 no es estacionaria.

Corregimos la no estacionariedad de la serie diferenciando la misma, obteniendo así la rentabilidad de cada índice.

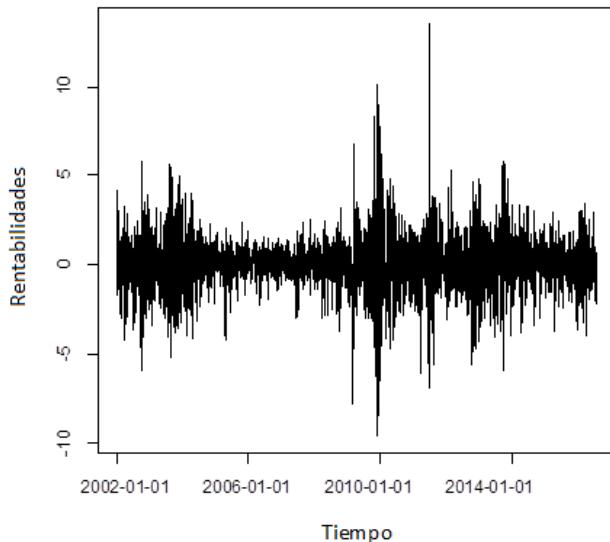


Figura 4-3: Gráfico de rentabilidades contra el tiempo

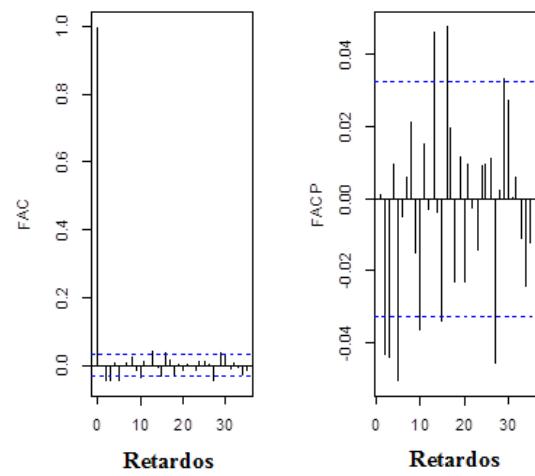


Figura 4-4: Correlograma rentabilidades

Gráficamente podemos observar que la no estacionariedad se ha corregido (Figura 4-3), la serie de rentabilidades es una serie estacionaria. El correlograma (Figura 4-4) corrobora la estacionariedad de la serie de rentabilidades. Además, realizando el mismo test de raíces unitarias, aplicado anteriormente al logaritmo de la serie de precios, obtenemos un *p*.valor de 0.01 para el caso del contraste Dickey-Fuller, que muestra rechazo a la hipótesis nula de existencia de raíces unitarias para un nivel de significatividad del 5%. Y un *p*.valor de 0.0799, en el caso del KPSS, para los niveles de significatividad del 1% y 5% presenta evidencia a favor de la hipótesis nula de estacionariedad. Así pues podemos afirmar que la serie de rentabilidades es estacionaria.

En la Tabla 4.1. y la Figura 4-5 se muestran los resultados obtenidos al realizar un análisis exploratorio, tanto gráfico como numérico, de la serie de rentabilidades del IBEX35. En dicho análisis se han aplicado los contrastes de media nula de *t* de Student, mediana nula de Wilcoxon, test de asimetría de Fisher, test de curtosis de Fisher, test de normalidad de Jarque-Bera (J-B) y de Shapiro-Wilks (S-W). Se observa que la rentabilidad media no es significativamente distinta de 0 pero la mediana sí lo es reflejando el que en el período analizado el IBEX35 tuvo más períodos de subida que de bajada aunque su impacto en la rentabilidad media no fue significativo. La desviación típica es significativamente más alta lo cual refleja que las oscilaciones de la rentabilidad fueron bastante más altas que sus valores medios y, por tanto, que los niveles de riesgo asociados al IBEX35 fueron, en general, altos. Observando, además, la

evolución diaria de las rentabilidades (Figura 4-3) se aprecia claramente el carácter heteroscedástico de la misma así como la existencia de *volatility clustering* (es decir, agrupamiento de la volatilidad en bloques), existiendo periodos en los que la volatilidad es menor (que coinciden con periodos en los que los rendimientos no sufren grandes cambios) y periodos en los que la volatilidad es mayor (momentos en los que la variación de los rendimientos respecto de su media es mayor).

Existe además una clara falta de normalidad en la serie, rechazada por los contrastes JB como SW. Ello es debido a la existencia de una clara leptocurtosis (ver histograma y QQ plot en la Figura 4-5) con un coeficiente de curtosis elevado (5,206) y una ligera asimetría positiva (0,0996 aunque ésta última se debe a un atípico positivo muy fuerte), siendo todas ellas significativas (los resultados de los test de asimetría y curtosis son significativos). Ello es debido a la existencia de periodos de tiempo con elevados valores de la rentabilidad en valor absoluto (los valores mínimo y máximo de la rentabilidad diaria fueron de un -9.58% y un 13.48%, respectivamente) junto a la existencia de otros periodos en los que la rentabilidad fue esencialmente nula. Todo ello pone de manifiesto los altos niveles de riesgo asociados a la evolución diaria de la serie.

Tabla 4.1. Análisis exploratorio de las rentabilidades de IBEX35

Media	Desv. Típica	Mediana	Mínimo	Máximo	Asimetría	Curtosis	J-B	S-W
0.0064	1.5153	0.0706 **	-9.5859	13.4836	0.0996 *	5.206 **	4128.28 **	0.95 **

* (**) significativo al 5% (1%)

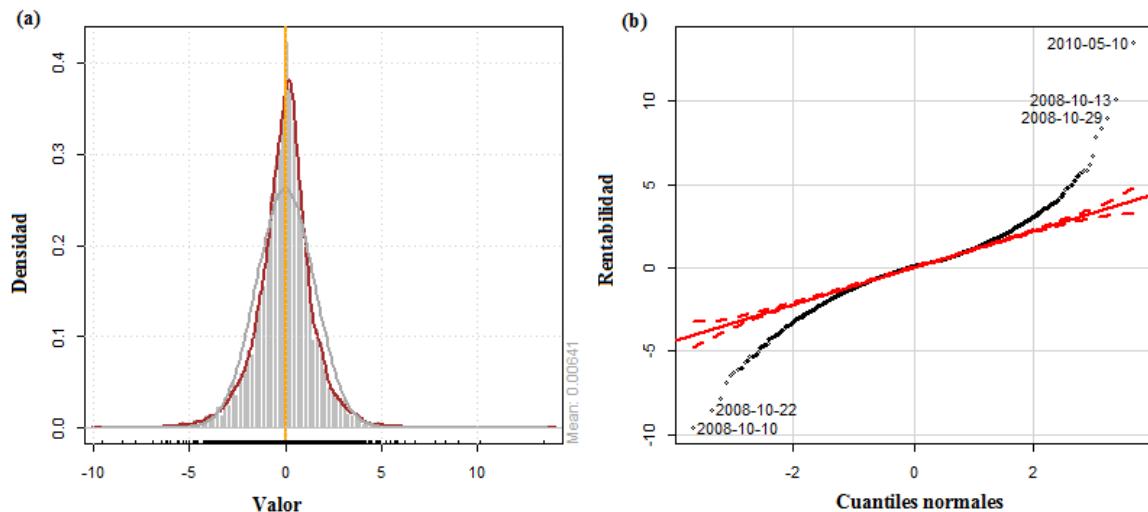


Figura 4-5: Gráfico de densidad (a) y gráfico cuantil-cuantil (b) de la serie de rentabilidades

Si analizamos las dependencias en media (ver correlogramas de la Figura 4-4) no se aprecia una estructura clara con autocorrelaciones significativas en los retardos 1, 2, 3 y 5. Por dicha razón recurrimos a criterios de selección de modelos y, más concretamente al criterio BIC de Schwarz (1976) que es consistente. En la Figura 4-6 se muestran los valores obtenidos para modelos ARMA(r,s) con $0 \leq r \leq 2$, $0 \leq s \leq 2$.

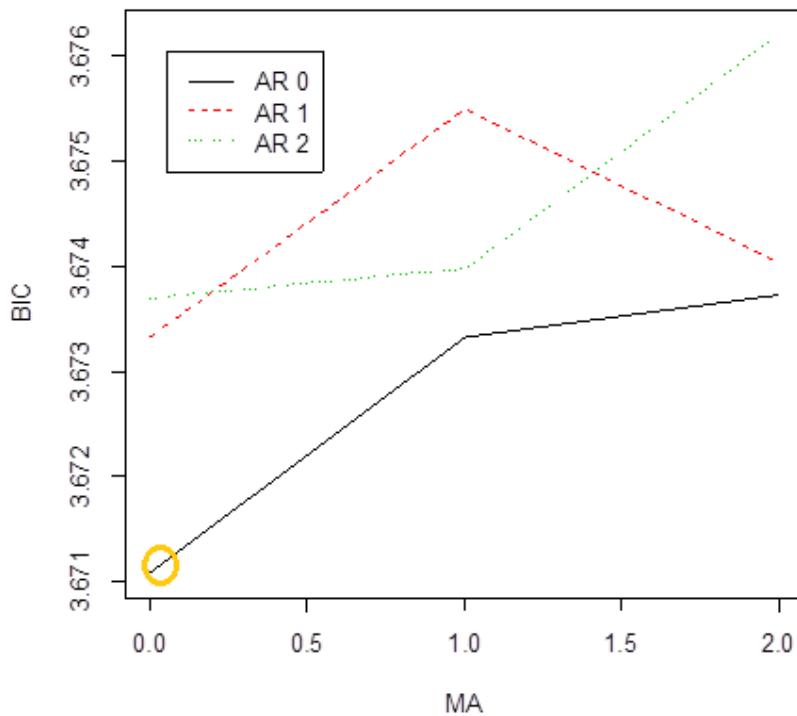


Figura 4-6: Gráfico BIC

El criterio a tener en cuenta para seleccionar un modelo es que el valor de BIC debe ser el menor de todos. Se observa (Figura 4-6) que el menor valor BIC, señalado con un círculo amarillo corresponde a un ARMA (0,0), también conocido como ruido blanco no apreciándose, por tanto, la existencia de dependencias en media relevantes.

Una vez identificado el orden del proceso ARMA que se ajusta al IBEX35, vamos a determinar qué modelo es el más adecuado para describir la evolución de la volatilidad. Dado que la serie parece ser ruido blanco, analizaremos los cuadrados de la misma, sus correlogramas y aplicaremos los contrastes ARCH y de Box-Ljung con el fin de analizar si la serie es homoscedástica. En la Tabla 4.3 y las Figuras 4-7 y 4-8 se muestran los resultados obtenidos. Se observa que todos los contrastes rechazan la hipótesis de homoscedasticidad, existiendo dependencias significativas positivas en todos los retardos en el correlograma de la Figura 4-8. Este hecho pone de manifiesto, de nuevo, la existencia de *volatility clustering* con una alta persistencia de la volatilidad a lo largo del tiempo, reflejando la tendencia a de la serie a tener largos períodos de altas (bajas) oscilaciones como consecuencia de dicha persistencia.

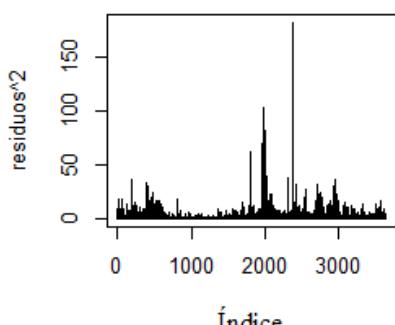


Figura 4-7: Gráfico de la serie de residuos al cuadrado

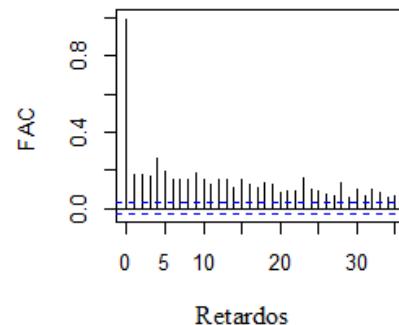


Figura 4-8: Correlograma de la serie de residuos al cuadrado

Dado que hay valores significativos tanto en las autocorrelaciones como en las autocorrelaciones parciales, ajustamos un modelo GARCH a la serie analizando el comportamiento de los residuos tipificados. Realizado este proceso para diversos procesos GARCH (cuyos resultados omitimos por brevedad) ajustamos un modelo GARCH (4,1) que es el que muestra una menor autocorrelación en los cuadrados de los residuos. En la Tabla 4.2 se muestran las estimaciones obtenidas para los modelos GARCH con mejores niveles de ajuste tanto para innovaciones normales como para innovaciones no normales. Estos últimos fueron seleccionados considerando las distribuciones descritas en la sección 2.1 para las innovaciones, analizando su bondad

de ajuste y eligiendo el mejor modelo de acuerdo al criterio BIC. Se aprecia que los modelos GARCH presentan menores valores de dicho criterio que los modelos homoscedásticos, siendo significativos tanto los parámetros α como β de dichos modelos y observándose altos niveles de persistencia (superiores a 0.99) en todos ellos fiel reflejo del elevado grado de *volatility clustering* presente en la serie.

Finalmente se observa que las estimaciones de los parámetros v y ξ de los modelos con innovaciones t de Student asimétricas ponen de manifiesto la existencia de leptocurtosis condicional (valores de v pequeños) y asimetría negativa ($\xi < 1$) significativas, si bien ésta última no muy fuerte, corroborando la falta de normalidad de los residuos tipificados apreciada en los resultados mostrados en la Tabla 4.3 y en la Figura 4-12.

Tabla 4.2: Resultados de la estimación de los modelos ajustados a las rentabilidades del IBEX35

	ARMA(0,0)		ARMA(0,0)-GARCH(4,1)		ARMA(0,0)-APARCH(1,1)	
	Normal	Exponencial Generalizada Asimétrica	Normal	t de Student asimétrica	Normal	t de Student asimétrica
$\hat{\mu}$	0.0064	0.0064	0.0640 **	0.0558 **		
$\hat{\omega}$	2.295	2.295	0.0254 **	0.0132 **	0.0237 **	0.0205 **
$\hat{\alpha}_1$			0.0618 **	0.0610 **	0.0620 **	0.0609 **
$\hat{\alpha}_2$			0.0325	0.0234		
$\hat{\alpha}_3$			0.0126	0.0014		
$\hat{\alpha}_4$			0.0000	0.0000		
$\hat{\beta}_1$			0.8844 **	0.9108 **	0.9318 **	0.9351 **
Persistencia $\hat{\alpha}(1)+\hat{\beta}(1)$	0.000	0.000	0.9913	0.9966		
$\hat{\gamma}_1$					1 **	1 **
$\hat{\delta}$					1.0732 **	1.0892 **
\hat{v}				9 **		10 **
$\hat{\xi}$				0.9166 **		0.8919 **
BIC	3.669	3.519	3.353	3.327	3.299	3.283

*(**) significativo al 5% (1%)

En la Tabla 4.3 y en las Figuras 4-9 a 4-12 se muestran los resultados del análisis de la bondad de ajuste de los modelos estimados a los datos. El estudio de la bondad de

ajuste de un modelo GARCH se basa en el análisis de los residuos tipificados $\hat{z}_t = \frac{\hat{\epsilon}_t}{\sqrt{\hat{\sigma}_t^2}}$

donde $\hat{\epsilon}_t$ son los residuos estimados y $\hat{\sigma}_t^2$ son las volatilidades estimadas, por el modelo ajustado. Si el modelo ajustado es el verdadero, dichos residuos serán ruido blanco débil homoscedástico. En las Figuras 4-9 y 4-10 se analizan las dependencias en media de los residuos correspondientes al modelo ARMA(0,0)-GARCH(4,1) con errores normales y errores t de Student asimétricas. No se aprecia la existencia de dependencias significativas en media.

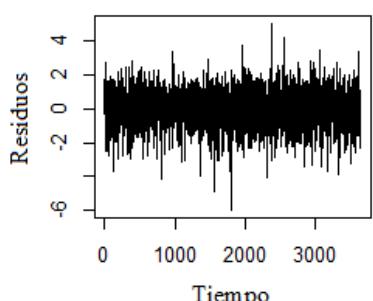


Figura 4-9: Gráfico de los residuos contra el tiempo

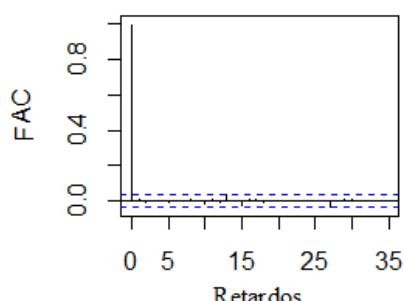
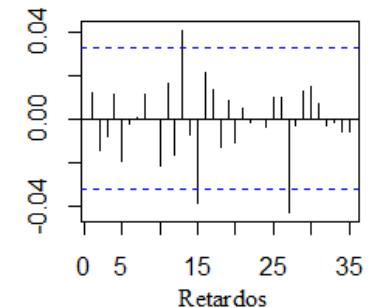


Figura 4-10: Correlograma de los residuos



En la Figuras 4-11 y 4.12 se analizan los residuos cuadráticos. Se observa la existencia de una dependencia en el retardo 4 significativa pero no importante siendo los demás retardos no significativos. Finalmente en la Figura 4.12 se muestran los ajustes de la distribución residual a los residuos tipificados mediante QQ plots mientras que la Tabla 4.3 analiza la bondad de ajuste en términos de los contrastes de hipótesis de Jarque-Bera y Kolmogorov Smirnov. Se rechaza la hipótesis de normalidad pero se acepta, débilmente al 5%, la hipótesis de que la distribución de los errores es una t de Student asimétrica.

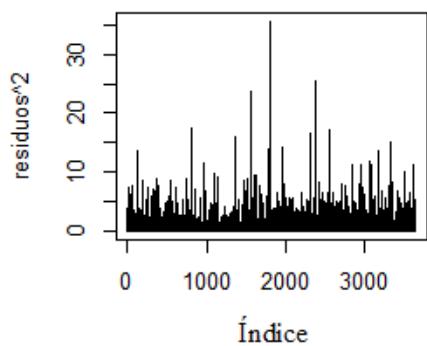


Figura 4-11: Gráfico de la serie de la serie de residuos al cuadrado

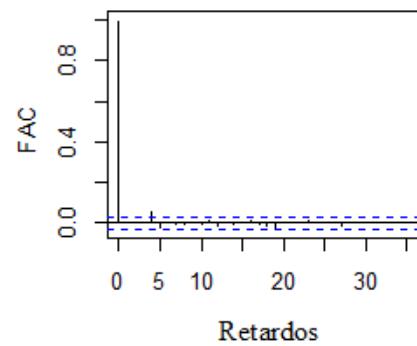
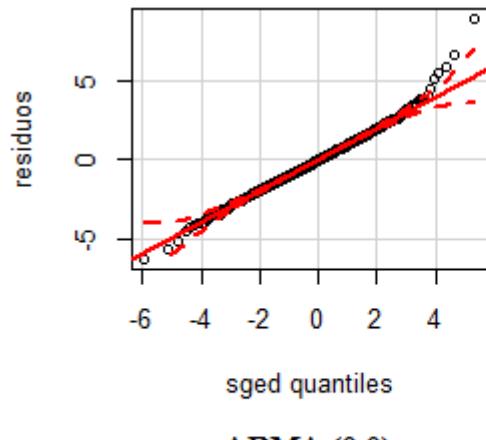
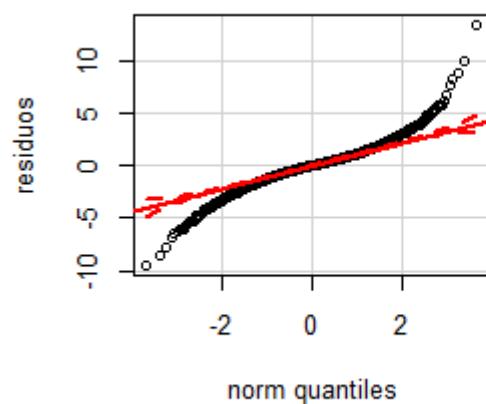
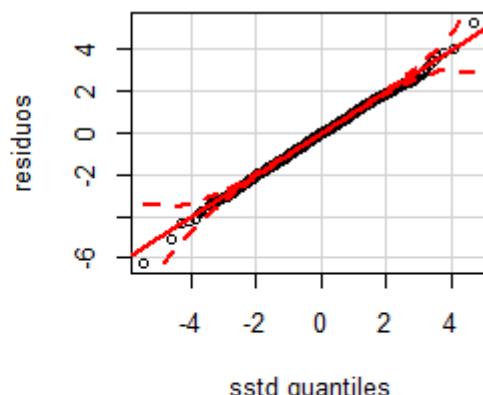
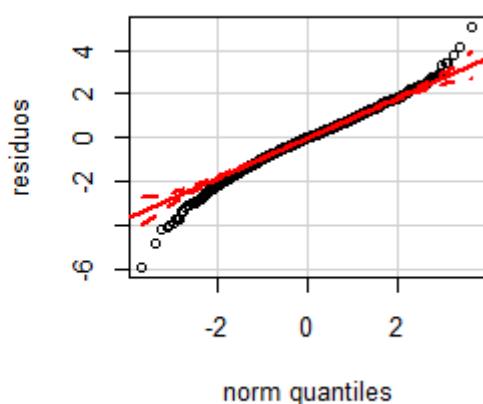


Figura 4-12: Correlograma de la serie de residuos al cuadrado



ARMA (0,0)



ARMA (0,0)-GARCH (4,1)

Figura 4-13: QQ plots de los residuos tipificados

Tabla 4.3: Bondad de ajuste de los modelos estimados ARMA y GARCH para las rentabilidades del IBEX35

	ARMA(0,0)		ARMA(0,0)-GARCH(4,1)	
	Normal	Exponencial Generalizada Asimétrica	Normal	t de Student asimétrica
Box-Ljung Residuos (8 retardos)	25.4012 **	25.4012 **	4.4009	3.8535
Box-Ljung Residuos² (8 retardos)	1049.69 **	1049.69 **	18.1319 *	23.507 **
Test ARCH (10 retardos)	6.262 **	6.262 **	23.1879 *	29.3125 **
Jarque-Bera	4128.28 **	4128.28 **	266.440 **	337.0544 **
Test K-S	0.0587 **	0.0118	0.0346 **	0.0246 *

* (**) significativo al 5% (1%)

Finalmente, y con el fin de analizar la posible existencia de efecto asimétrico, en la Figura 4-14 se muestran las correlaciones cruzadas de los residuos tipificados con sus cuadrados apreciándose la existencia de correlaciones significativamente negativas en los retardos negativos. Existe, por tanto, evidencia de una relación inversa de los residuos retardados con sus cuadrados (estimaciones de la volatilidad) y, por tanto, de la existencia de un efecto asimétrico del tipo predicho por la teoría financiera.

Para analizar la significación de dicho efecto estimamos un modelo APARCH (1,1) con errores normales y no normales, en este caso distribuidos según una t de Student asimétrica. Los resultados de la estimación se muestran en la Tabla 4.2. Se observa que el parámetro γ es significativamente positivo en ambos modelos por lo cual existe evidencia de la existencia de un efecto asimétrico significativo del mismo tipo que el pronosticado por la Teoría Financiera. Además el ajuste del modelo a los datos es mejor que el del modelo GARCH dando lugar a un menor valor del criterio BIC. Los valores estimados del parámetro δ son significativamente distintos de 2 pero no de 1 por lo que, en ambos modelos, tienden a describir la evolución de la volatilidad medida en términos de la desviación típica y no de la varianza. Finalmente, si analizamos la bondad de ajuste (ver Tabla 4.4 y Figuras 4-15 a 4-19) observamos que los residuos tipificados son ruido blanco homoscedástico pero no es normal, siendo la distribución

t de Student asimétrica la que mejor se ajusta a su evolución. Así mismo una vez estimado dicho modelo prácticamente desaparece la significación de las correlaciones cruzadas entre los residuos y sus cuadrados (ver Figura 4-14) por lo que podemos concluir que, de los modelos considerados en el trabajo es el que mejor se ajusta a la evolución de las rentabilidades del IBEX35.

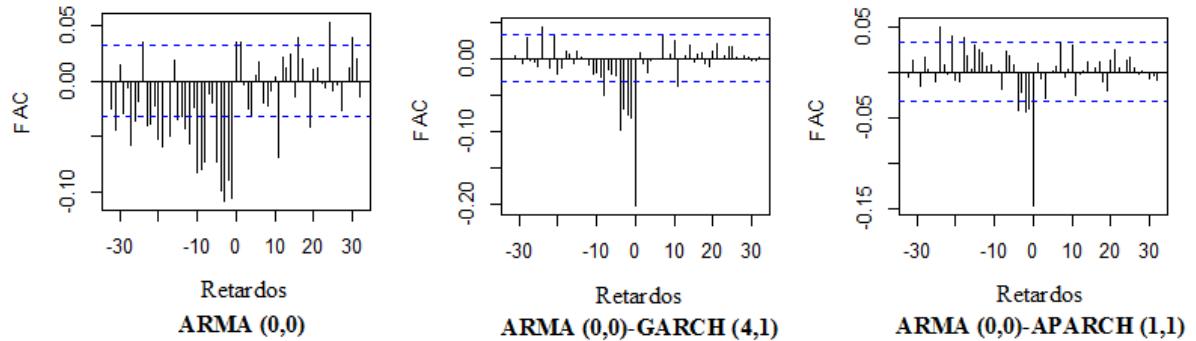


Figura 4-14: Correlaciones cruzadas de residuos y residuos cuadráticos

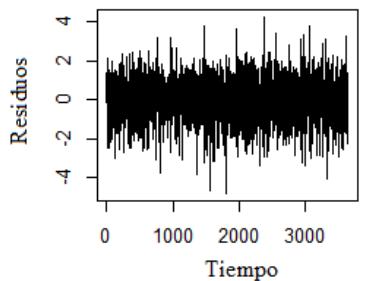


Figura 4-15: Gráfico de los residuos contra el tiempo

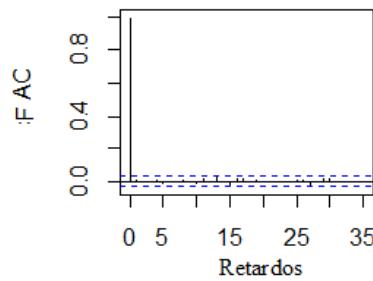


Figura 4-16: Correlograma de los residuos

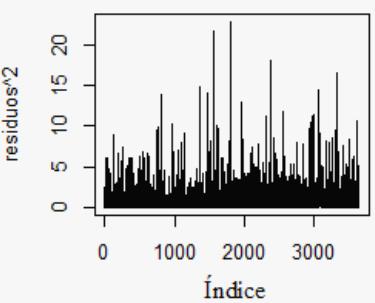


Figura 4-17: Gráfico de la serie de residuos al cuadrado

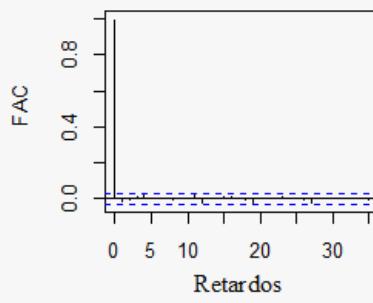
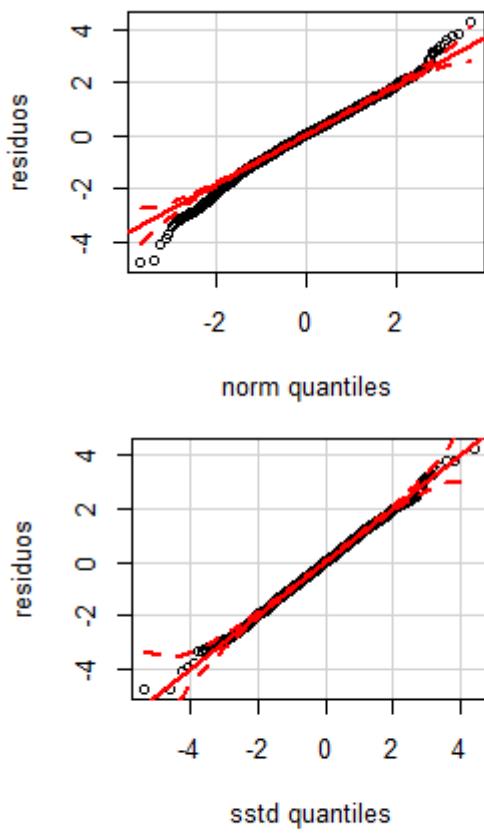


Figura 4-18: Correlograma de la serie de residuos al cuadrado



ARMA (0,0)-APARCH (1,1)

Figura 4-19: QQ plots de los residuos tipificados

Tabla 4.4: Bondad de ajuste de los modelos estimados APARCH para las rentabilidades del IBEX35

	ARMA(0,0)- APARCH(1,1)	
	Normal	t de Student asimétrica
Jarque-Bera	131.183 **	142.671 **
Test K-S	0.0318 **	0.0124
Box-Ljung Residuos (8 retardos)	3.8563	3.7722
Box-Ljung Residuos² (8 retardos)	11.6349	10.9124
Test ARCH (10 retardos)	17.6493	16.9505

* (**) significativo al 5% (1%)

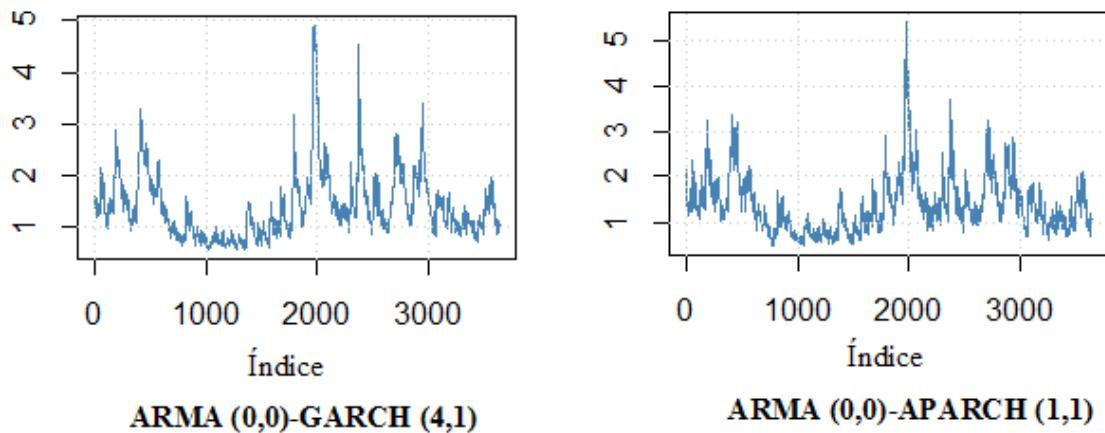


Figura 4-20: Gráfico de volatilidad estimada

Finalmente en la Figura 4-20 se muestran las volatilidades estimadas por los modelos ARMA (0,0)-GARCH (4,1) y ARMA (0,0)-APARCH (1,1) con errores no normales. Se aprecia que dichas volatilidades son muy similares observándose la existencia de altos valores de la volatilidad en el periodo 0-800 (02/01/2001 a 13/03/2003), un descenso con niveles bajos de volatilidad en el periodo 800-1700 (13/03/2003 a 30/08/2005) y un repunte en el periodo 1700 hasta el final (30/08/2005 a 24/04/2015), podemos observar que los periodos en los que la volatilidad crece coinciden con el periodo de crisis financiera en España, la crisis del 2001 y la crisis del 2007, especialmente la crisis de 2007 que fue una crisis de carácter financiero afectó en mayor medida a la volatilidad, pues está reflejó con una fuerte subida los cambios producidos en el precio de dicho activo.

4.2. Resultados del proceso de estimación y selección de modelos para el resto de los índices

En los anexos se muestran los resultados obtenidos en el proceso de estimación y selección de modelos para el resto de los índices bursátiles así como el análisis de la bondad de ajuste de cada uno de ellos.

Como ocurría con el IBEX35, tenemos que la serie de precios de cada uno de los demás índices es no estacionaria, como podemos observar gráficamente a través de los gráficos de precios contra el tiempo y sus respectivos correlogramas (ver gráfico en el punto 1.a), así como analíticamente a través de los contrastes de raíces unitarias de

Dickey-Fuller y KPSS (ver tabla en el punto 2). Por lo que convertiremos la serie de precios en serie de rentabilidades a través de la variación de sus logaritmos.

La serie de rentabilidades sí es estacionaria, lo hemos comprobado utilizando el mismo procedimiento que para la serie de precios (ver gráfico en el punto 1.b y tabla en el punto 2). A través de un análisis exploratorio inicial (ver tabla en el punto 3) tenemos constancia de que para ningún índice la media es significativa, mientras que la mediana sí, lo que implica que tuvieron más periodos de subida que de bajada aunque su impacto en la rentabilidad media no fue significativo. Salvo para el caso del índice EUROSTOXX, que se rechaza que la mediana sea significativa, es decir, que tuvo más periodos de bajada que de subida. Como ocurría con el IBEX35, la desviación típica es significativamente más alta lo cual refleja que las oscilaciones de la rentabilidad fueron bastante más altas que sus valores medios y, por tanto, que los niveles de riesgo asociados a los índices fueron, en general, altos. Observando, además, la evolución diaria de las rentabilidades (ver gráfico en el punto 1.b) se aprecia claramente el carácter heteroscedástico de la misma así como la existencia de *volatility clustering* (es decir, agrupamiento de la volatilidad en bloques), existiendo periodos en los que la volatilidad es menor (que coinciden con periodos en los que los rendimientos no sufren grandes cambios) y periodos en los que la volatilidad es mayor (momentos en los que la variación de los rendimientos respecto de su media es mayor). En términos de normalidad, mediante el test de curtosis de Fisher y los contrastes de Jarque-Bera y Shapiro-Wilks, tenemos que en las cinco series se rechaza la hipótesis de normalidad, como ocurría con el IBEX35, ya que habíamos comentado que es un rasgo característico de cualquier serie financiera. También podemos observar la existencia de leptocurtosis (curtosis elevada) y falta de normalidad a través de los gráficos cuantil-cuantil (ver histograma y QQ plot en el punto 4). Por último, respecto al análisis exploratorio, la hipótesis de simetría no se rechaza salvo en el caso del índice NIKKEI en el que se aprecia una ligera asimetría negativa debido a la presencia de algunos atípicos negativos.

En las tablas del punto 5 de los anexos se muestran las estimaciones obtenidas para los modelos con mejores niveles de ajuste tanto para innovaciones normales como para innovaciones no normales. Estos últimos fueron seleccionados considerando las distribuciones descritas en la sección 2.1 para las innovaciones, analizando su bondad

de ajuste y eligiendo el mejor modelo de acuerdo al criterio BIC. Se aprecia que los modelos GARCH presentan menores valores de dicho criterio que los modelos heteroscedásticos, siendo significativos tanto los parámetros α como β de dichos modelos y observándose altos niveles de persistencia (superiores a 0.99) en todos ellos fiel reflejo del elevado grado de *volatility clustering* presente en la serie.

Finalmente se observa que las estimaciones de los parámetros v y ξ de los modelos con innovaciones t de Student asimétrica ponen de manifiesto la existencia de leptocurtosis condicional (valores de v pequeños) y ligera asimetría negativa ($\xi < 1$ con valores en torno a 0.9) significativas.

Pasamos al análisis de los residuos de las series de rentabilidades mediante el estudio de la bondad de ajuste (ver tabla en el punto 6 y gráficos en los puntos 7, 8 y 10). Vamos a estudiar los tres problemas mencionados anteriormente en el estudio del IBEX35: falta de normalidad (test Jarque-Bera y prueba de Kolmogorov-Smirnov), autocorrelación (contraste Box-Ljung) y heteroscedasticidad (test ARCH). En primer lugar, vemos que el problema de falta de normalidad continúa en la serie de residuos de las rentabilidades para cada uno de los índices. En cuanto a la autocorrelación, tenemos que para los índices BEL20, DAX y DOWJONES, la autocorrelación de los residuos así como la de los residuos cuadráticos se corrige al introducir en el modelo inicial tanto la parte GARCH como la APARCH. Para el índice EUROSTOXX también existen problemas de autocorrelación en el modelo ARMA que se solucionan al introducir los modelos GARCH y APARCH, pero con la salvedad de que en este caso el modelo GARCH donde se distribuyen los residuos según una t de Student asimétrica, existen problemas de autocorrelación al 5%. Y ya para acabar con el análisis de autocorrelación de los residuos y los residuos al cuadrado, el índice NIKKEI no muestra problemas de autocorrelación en los residuos pero sí en los residuos cuadráticos, que se corrige al incorporar los modelos GARCH y APARCH al modelo inicial. Por último, el problema de heteroscedasticidad, medido a través del test ARCH, el problema de que la varianza no sea constante se corrige para todos los índices incorporando al modelo las partes GARCH y APARCH, sin embargo, para el índice BEL20, la heteroscedasticidad continúa en el modelo GARCH. Atendiendo al criterio BIC, los modelos elegidos son los siguientes: para BEL20, APARCH-t de Student asimétrica (2.872); para DAX, APARCH-exponencial generalizada asimétrica (3.265); para DOWJONES, APARCH-

exponencial generalizada asimétrica (2.659); para EUROSTOXX, APARCH-t de Student asimétrica (3.246) y; para NIKKEI, APARCH-exponencial generalizada asimétrica (3.446). En todos los casos se rechaza, por tanto, la hipótesis de normalidad para la distribución condicional del error, siendo todas las distribuciones leptocúrticas y con asimetría negativa significativa pero no muy fuerte.

Si se analiza el gráfico 9 de las correlaciones cruzadas de los residuos cuadráticos con los residuos retardados se aprecia la existencia de correlaciones significativa de signo negativo en los retardos negativos señalando la posible presencia de efecto asimétrico en la serie. La introducción del modelo APARCH mejora, en general, el ajuste del modelo a los datos medido a través del criterio BIC y es capaz de corregir problemas de autocorrelación y heteroscedasticidad, como hemos visto en el estudio de la bondad de ajuste de los residuos de la serie de rentabilidades. Analizando las estimaciones de dicho modelo se observa que los coeficientes γ son, en general, positivos, lo cual concuerda con lo predicho por la teoría financiera acerca de la existencia de efecto asimétrico: existe e implica que el impacto ejercido sobre la volatilidad por una sobrevaloración del mercado (residuo negativo) es superior al ejercido por una infravaloración.

Finalmente en el gráfico 11 se muestran las volatilidades estimadas por los modelos ARMA-GARCH y ARMA-APARCH con errores no normales. Se aprecia que dichas volatilidades son muy similares observándose la existencia de altos valores de la volatilidad en el periodo 0-800 (02/01/2001 a 13/03/2003), un descenso con niveles bajos de volatilidad en el periodo 800-1700 (13/03/2003 a 30/08/2005) y un repunte en el periodo 1700 hasta el final (30/08/2005 a 24/04/2015).

5. Validación predictiva extramuestral de los modelos estimados

La evaluación de los pronósticos es un paso clave en cualquier ejercicio de esta naturaleza. Sin embargo, el estudio de la calidad de los diferentes modelos de volatilidad puede ser muy complejo, debido a que no existe un único criterio capaz de seleccionar el mejor modelo. Debido a ello, muchos autores han destacado la importancia de evaluar los pronósticos de la volatilidad por medio de funciones de pérdida reales que enfrenta el usuario final, proponiendo criterios basados en los beneficios o en la utilidad. Desafortunadamente, no podemos contar con esa

información, pues las funciones de pérdida dependen de las preferencias desconocidas y no observables de los agentes económicos.

En esta sección se realiza un estudio comparativo del comportamiento predictivo extramuestral de los modelos estimados en la sección anterior. Para ello realizaremos una validación predictiva tipo “*rolling*” de cada uno de ellos en el que, para cada periodo de tiempo t , estimaremos los parámetros del modelo utilizando la información disponible hasta el periodo $t-1$ y luego realizaremos una predicción de la rentabilidad de la serie (o del precio según el criterio utilizado) en el periodo t . La validación de la predicción se hará en términos puntuales analizando el comportamiento del error de predicción a 1 paso vista $\hat{e}_t = r_t - E[r_t | \hat{\zeta}_{1:(t-1)}, \hat{\psi}_{1:(t-1)}, \Omega_{t-1}]$ donde $\hat{\zeta}_{1:(t-1)}, \hat{\psi}_{1:(t-1)}$ son los valores estimados de los parámetros de la media y la varianza utilizando la información Ω_{t-1} , y en términos de intervalos de predicción a 1 paso vista de un nivel de confianza $1-\alpha$ con $0 < \alpha < 1$, donde analizaremos el comportamiento predictivo de los intervalos de predicción:

$$IPRED_{1-\alpha}(r_t) = (\hat{r}_{t,inf}(1-\alpha), \hat{r}_{t,sup}(1-\alpha)) =$$

$$\left(E[r_t | \hat{\zeta}_{1:(t-1)}, \hat{\psi}_{1:(t-1)}, \Omega_{t-1}] - d_{\frac{\alpha}{2}} D[r_t | \hat{\zeta}_{1:(t-1)}, \hat{\psi}_{1:(t-1)}, \Omega_{t-1}], E[r_t | \hat{\zeta}_{1:(t-1)}, \hat{\psi}_{1:(t-1)}, \Omega_{t-1}] + d_{\frac{\alpha}{2}} D[r_t | \hat{\zeta}_{1:(t-1)}, \hat{\psi}_{1:(t-1)}, \Omega_{t-1}] \right)$$

donde

$E[r_t | \hat{\zeta}_{1:(t-1)}, \hat{\psi}_{1:(t-1)}, \Omega_{t-1}]$ y $D[r_t | \hat{\zeta}_{1:(t-1)}, \hat{\psi}_{1:(t-1)}, \Omega_{t-1}]$ denotan la media y desviación típica condicionales y d_α denota el cuantil $1-\alpha$ de la distribución condicional del error tipificado z_t .

En lo que sigue comenzaremos describiendo cómo se calculan $E[r_t | \zeta, \psi, \Omega_{t-1}]$ y $D[r_t | \zeta, \psi, \Omega_{t-1}]$ para los modelos considerados en el trabajo; posteriormente se mostrarán los criterios utilizados para evaluar las predicciones realizadas para, finalmente, presentar los resultados obtenidos.

5.1. Predicción en modelos ARMA-GARCH

Todos los modelos ajustados son casos particulares del modelo

$$r_t = G(x_t; \zeta) + \sigma_t(x_t; \psi) z_t$$

$$z_t | \Omega_{t-1} \sim D(0, 1, \lambda)$$

donde x_t son las variables explicativas del modelo que, en nuestro caso son valores pasados de la serie de los residuos o de los residuos cuadráticos. Para estos casos se verifica que:

$$E[r_t | \zeta, \psi, \lambda, \Omega_{t-1}] = G(x_t; \zeta) \text{ donde } \Omega_t = \{r_1, \dots, r_t, x_1, \dots, x_{t+1}\}$$

$$\text{Var}[r_t | \zeta, \psi, \lambda, \Omega_{t-1}] = \sigma_t^2(x_t, \psi)$$

y $d_\alpha(\lambda)$ será el cuantil $1-\alpha$ de la distribución $D(0, 1, \lambda)$.

5.2. Criterios de comparación

En esta sección describimos, brevemente, los criterios empleados en el proceso de validación predictiva de los modelos. Dichos criterios se pueden agrupar en dos grupos: aquéllos que evalúan su comportamiento predictivo puntual y aquéllos que evalúan su comportamiento predictivo en términos de intervalos.

5.2.1. Evaluación del comportamiento predictivo puntual

Utilizamos dos criterios que evalúan el comportamiento del error predictivo de la rentabilidad $r_t - \hat{r}_t$: la raíz del error cuadrático medio (RMSE) y el error absoluto medio (MAD) cuyas expresiones vienen dadas por:

$$\text{RMSE} = \sqrt{\sum_{t=t_0}^T \frac{(r_t - \hat{r}_t)^2}{T - t_0 + 1}}$$

$$\text{MAD} = \sum_{t=t_0}^T \frac{|r_t - \hat{r}_t|}{T - t_0 + 1}$$

donde $\hat{r}_t = E[r_t | \hat{\zeta}_{1:(t-1)}, \hat{\psi}_{1:(t-1)}, \Omega_{t-1}]$. También utilizamos el coeficiente de desigualdad de Theil que viene dado por:

$$\text{THEIL} = \frac{\sqrt{\sum_{t=t_0}^T (r_t - \hat{r}_t)^2}}{\sqrt{\sum_{t=t_0}^T \hat{r}_t^2} + \sqrt{\sum_{t=t_0}^T r_t^2}}$$

que evalúa dicho error en términos relativos comparando el valor predictivo del modelo con el ruido blanco. Este coeficiente toma valores entre 0 y 1 de forma que cuánto más (menos) cercano esté su valor a 1 peor (mejor) es su comportamiento con respecto al modelo de ruido blanco. Finalmente también utilizamos el error absoluto medio de predicción del precio medio en términos relativos PMAD que viene dado por:

$$PMAD = 100 \frac{1}{T-t_0+1} \sum_{t=t_0}^T \frac{|P_t - \hat{P}_t|}{|P_t|}$$

donde $\hat{P}_t = \exp[\hat{r}_t]$. En los 4 casos t_0 denota el inicio del periodo de validación predictiva extramuestral dejando los t_{0-1} primeros periodos de tiempo para obtener la estimación inicial del modelo.

Un segundo subgrupo de criterios se basa en la descomposición del error cuadrático medio de predicción en la suma de 3 términos: el sesgo, la varianza y la covarianza, analizando qué proporción de dicho error corresponde a cada término.

Dichas proporciones vienen dadas por:

- Proporción del sesgo:

$$\frac{\left(\sum_{t=t_0}^T \frac{\hat{r}_t}{T-t_0+1} - \sum_{t=t_0}^T \frac{r_t}{T-t_0+1} \right)^2}{\sum_{t=t_0}^T \frac{(r_t - \hat{r}_t)^2}{T-t_0+1}}$$

que mide la cuantía de posibles errores sistemáticos en la predicción.

- Proporción de la varianza:

$$\frac{(s_{\hat{r}} - s_r)^2}{\sum_{t=t_0}^T \frac{(r_t - \hat{r}_t)^2}{T-t_0+1}}$$

donde $s_{\hat{r}}$ y s_r son las desviaciones típicas de $\{r_t; t = t_0, \dots, T\}$ y $\{\hat{r}_t; t = t_0, \dots, T\}$ y que mide si las predicciones puntuales del modelo capturan las oscilaciones de la serie

- Proporción de la covarianza:

$$\frac{2(1 - \text{Corr}(r_t, \hat{r}_t))s_r s_{\hat{r}}}{\sum_{t=t_0}^T \frac{(r_t - \hat{r}_t)^2}{T-t_0+1}}$$

donde $\text{Corr}(r_t, \hat{r}_t)$ es la correlación entre $\{r_t; t = t_0, \dots, T\}$ y $\{\hat{r}_t; t = t_0, \dots, T\}$, y que mide el grado de concordancia entre los valores predichos por el modelo y los observados.

5.2.2. Evaluación del comportamiento predictivo por intervalos

Los criterios utilizados evalúan, por un lado, el nivel de cubrimiento empírico de los intervalos de predicción y, por el otro, la precisión de la misma teniendo en cuenta, por un lado, la longitud de los intervalos de predicción y, por el otro, la magnitud de los errores de cubrimiento.

El nivel de cubrimiento empírico viene dado por la expresión:

$$\text{CUB}(1-\alpha) = \frac{100}{T-t_0+1} \sum_{t=t_0}^T I_{\text{IPRED}_{1-\alpha}(r_t)}(r_t)$$

donde I_A denota la función indicadora del conjunto A y $1-\alpha$ es el nivel de confianza. Además, para evaluar la significación estadística de dichos cubrimientos utilizamos los contrastes de Christoffersen (1998) que analizan los niveles de cubrimiento condicional e incondicional de los intervalos elaborados así como su independencia. Si $I_t = I_{\text{IPRED}_{1-\alpha}(r_t)}(r_t)$ se tiene que:

- Test LR de cubrimiento incondicional:

Contrasta si $E[I_t] = 1-\alpha$ utilizando el contraste LR que viene dado por

$$\begin{aligned} \text{LR}_{uc} &= -2 \log \left(\frac{L(\alpha | I_{t_0}, \dots, I_T)}{L(\hat{\alpha} | I_{t_0}, \dots, I_T)} \right) = -2 \log \left(\frac{\alpha^{n_1} (1-\alpha)^{n_0}}{\hat{\alpha}^{n_1} (1-\hat{\alpha})^{n_0}} \right) \text{ donde } \hat{\alpha} = \frac{n_1}{n_0 + n_1} \\ \text{pvalor}_{uc} &= P[\chi^2 \geq \text{LR}_{uc, \text{obs}}] \end{aligned}$$

donde $n_0 = \sum_{t=t_0}^T I_{\text{IPRED}_{1-\alpha}(r_t)}(r_t)$ es el número de aciertos y $n_1 = T - n_0$ es el número de fallos.

- Test LR de independencia:

Contrasta si $\{I_t; t=t_0, \dots, T\}$ son independientes

$$\begin{aligned} \text{Sean } n_{11} &= \sum_{t=t_0+1}^T I_{t-1} I_t, n_{10} = \sum_{t=t_0+1}^T I_{t-1} (1 - I_t) \\ n_{01} &= \sum_{t=t_0+1}^T (1 - I_{t-1}) I_t, n_{00} = \sum_{t=t_0+1}^T (1 - I_{t-1}) (1 - I_t) \end{aligned}$$

El estadístico del contraste es:

$$LR_{ind} = -2 \log \left(\frac{L(\hat{\alpha} | I_{t_0}, \dots, I_T)}{L(\hat{\pi}_{01}, \hat{\pi}_{11} | I_{t_0}, \dots, I_T)} \right) = -2 \log \left(\frac{(1-\hat{\alpha})^{n_{00}+n_{10}} \hat{\alpha}^{n_{01}+n_{11}}}{(1-\hat{\pi}_{01})^{n_{00}} \hat{\pi}_{01}^{n_{01}} (1-\hat{\pi}_{11})^{n_{10}} \hat{\pi}_{11}^{n_{11}}} \right)$$

$$pvalor_{uc} = P[\chi^2 \geq LR_{ind,obs}]$$

$$\text{donde } \hat{\pi}_{01} = \frac{n_{01}}{n_{00} + n_{01}}, \hat{\pi}_{11} = \frac{n_{11}}{n_{10} + n_{11}}, \hat{\alpha} = \frac{n_{01} + n_{11}}{n_{00} + n_{10} + n_{01} + n_{11}} = \frac{n_1}{n_0 + n_1}$$

- Test LR de cubrimiento condicional:

Contrasta si $E[I_t | I_{t-1}] = \alpha$

El estadístico del contrastes es:

$$LR_{cc} = LR_{uc} + LR_{ind} \text{ con } pvalor_{cc} = P[\chi^2 \geq LR_{cc,obs}]$$

Por su parte, para evaluar el nivel de exactitud del intervalo utilizamos la pérdida de Gneiting y Raftery (1007) que viene dada por:

$$\text{Loss}(1-\alpha) = 2 \sum_{t=t_0}^T (\hat{r}_{tinf}(\alpha) - r_t) I(r_t < \hat{r}_{tinf}(\alpha)) + 2 \sum_{t=t_0}^T (r_t - \hat{r}_{tsup}(\alpha)) I(r_t > \hat{r}_{tsup}(\alpha))$$

$$= 2 \sum_{t=t_0}^T (\hat{r}_{tsup}(\alpha) - \hat{r}_{tinf}(\alpha)) + \frac{2 \sum_{t=t_0}^T (\hat{r}_{tinf}(\alpha) - r_t) I(r_t < \hat{r}_{tinf}(\alpha)) + 2 \sum_{t=t_0}^T (r_t - \hat{r}_{tsup}(\alpha)) I(r_t > \hat{r}_{tsup}(\alpha))}{\alpha}$$

5.3. Resultados obtenidos

5.3.1. IBEX35

Para realizar la evaluación predictiva se tomó como periodo inicial $t_0=1316$ (11/08/2004). En la Tabla 5.1. se muestran los resultados obtenidos en la evaluación predictiva puntual mientras que las Tablas 5.2 y 5.3. contienen los resultados correspondientes a la evaluación predictiva por intervalos para los niveles de confianza del 95% y 99%, respectivamente. Finalmente la Figura 5.1 analiza los resultados gráficamente mostrando, la predicción puntual, los límites de los intervalos predictivos del 95% y 99% y los valores observados de la serie de rentabilidades para cada uno de los 6 modelos ajustados a la serie.

Se observa que, a nivel de rentabilidades, el comportamiento de los 6 modelos ajustados es muy similar, con diferencias existentes a nivel del tercer decimal. Ello es lógico debido a que, al ser todos los modelos ruido blanco, los valores de las medias de predicción coinciden con el valor estimado de μ que es muy similar para los 6 modelos.

Tabla 5.1: Resultados de la validación predictiva puntual

ARMA(0,0)		ARMA(0,0)-GARCH(4,1)		ARMA(0,0)-APARCH(1,1)	
	Normal		Normal		Normal
	Exponencial Generalizada Asimétrica		t de Student asimétrica		t de Student asimétrica
RMSE	1.6015	1.6015	1.6026	1.6025	1.6017
MAD	1.1283	1.1283	1.1274	1.1274	1.1277
PMAD	0.0351	0.0351	0.0337	0.0337	0.0542
Theil	0.9917	0.9917	0.9669	0.9681	0.9841
Sesgo	0.0026	0.0026	0.1059	0.0992	0.0218
Varianza	98.5276	98.5276	97.0338	97.1969	98.3068
Covarianza	1.4698	1.4698	2.8603	2.7039	1.6714
					1.6924

Dado que la predicción es esencialmente constante (ver Figura 5.1) la mayor proporción del error cuadrático medio corresponde a la varianza siendo los niveles de sesgo, en general, despreciables con los modelos homoscedásticos mostrando los niveles de sesgo más bajo debido a su menor grado de complejidad. Algo similar ocurre a nivel de predicción del precio en el que los modelos GARCH son los que mejor comportamiento muestran pero con diferencias poco relevantes con el resto de los modelos.

A nivel de intervalos, si analizamos los niveles de cubrimiento de los mismos todos los modelos tienden a infracubrir tanto al 95% como al 99% siendo los modelos con errores no normales los que mejor comportamiento tienen tanto a nivel incondicional como condicional. Las mayores diferencias, sin embargo, se dan a nivel de exactitud apreciándose que los modelos heteroscedásticos tienen niveles de exactitud significativamente más elevados que los homoscedásticos debido a su mayor adaptación a las fluctuaciones de la serie (ver Figura 5.1). La influencia del efecto asimétrico tiene un carácter más marginal, mejorando los niveles de exactitud pero en menor cuantía que la mejora lograda con respecto a los modelos homoscedásticos tanto a niveles del 95%

como del 99%. Finalmente el uso de distribuciones no normales no parece tener un efecto de mejora del nivel de exactitud siendo el modelo ARMA (0,0)-APARCH (1,1) con errores normales el que mejor comportamiento tiene a ambos niveles de confianza aunque con niveles de cubrimiento ligeramente inferiores.

Tabla 5.2: Resultados de la validación predictiva por intervalo al 95%

ARMA(0,0)		ARMA(0,0)-GARCH(4,1)		ARMA(0,0)-APARCH(1,1)	
Normal	Exponencial Generalizada Asimétrica	Normal	t de Student asimétrica	Normal	t de Student asimétrica
CUB(0.95)	93.1985%	94.3607%	93.371%	93.8872%	93.026%
pvalor LRu	0.0002	0.1656	0.0006	0.0172	0.0000
pvalor LRind	0.0000	0.0000	0.0699	0.0741	0.0048
pvalor LRc	0.0000	0.0000	0.0005	0.0119	0.0000
LOSS(0.95)	34786.74	35557.98	31947.08	32126.44	30902.04
					31018.22

Tabla 5.3: Resultados de la validación predictiva por intervalo al 99%

ARMA(0,0)		ARMA(0,0)-GARCH(4,1)		ARMA(0,0)-APARCH(1,1)	
Normal	Exponencial Generalizada Asimétrica	Normal	t de Student asimétrica	Normal	t de Student asimétrica
CUB(0.99)	97.1158%	98.6225%	97.675%	98.4933%	97.374%
pvalor LRu	0.0000	0.0837	0.0000	0.0224	0.0000
pvalor LRind	0.0000	0.0788	0.0087	0.0156	0.3088
pvalor LRc	0.0000	0.0478	0.0000	0.0039	0.0000
LOSS(0.99)	56393.98	56256.75	46351.58	47232.84	45748.86
					45909.59

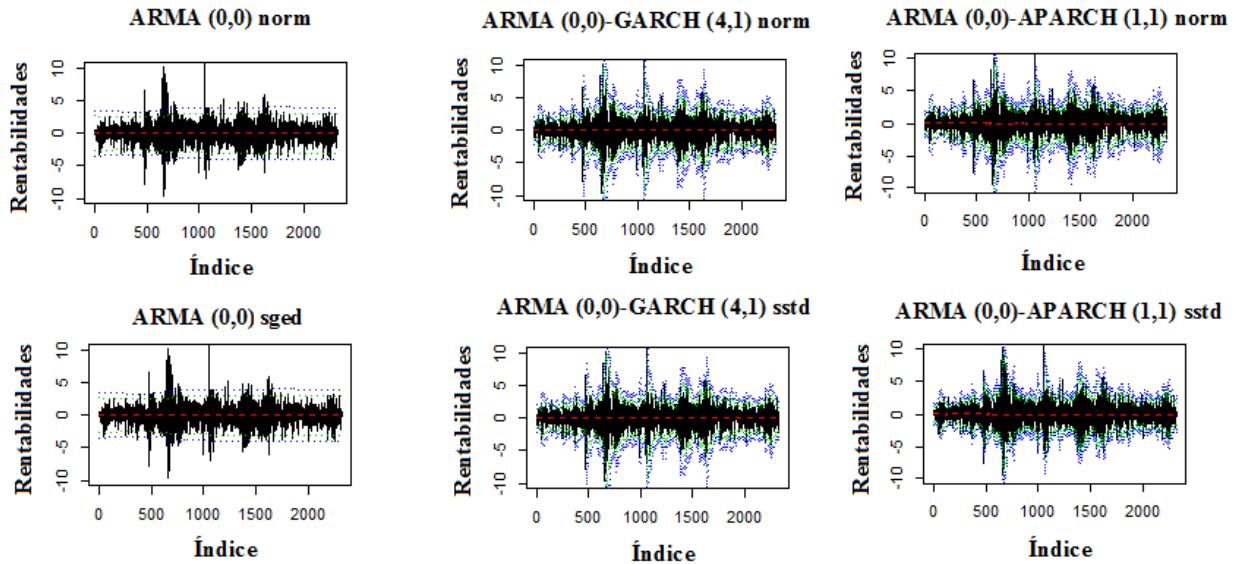


Figura 5-1: Resultados gráficos de la validación predictiva

5.3.2. Resto de los índices

Pasamos al último punto, analizando las predicciones, tanto puntuales como en intervalos realizadas por los diversos modelos ajustados a cada índice. Por un lado, las predicciones puntuales, muestran analíticamente que para los índices BEL20, EUROSTOXX y NIKKEI el modelo APARCH tendría menos error a la hora de predecir que el modelo GARCH, mientras que para DAX y DOWJONES el error en ambos modelos es prácticamente equivalente.

Respecto a la predicción por intervalos y, empezando con el índice BEL20, tenemos cobertura incondicional correcta para los modelos ARMA distribuido según una exponencial generalizada asimétrica y para el GARCH con una t de Student asimétrica, en este caso sólo al 1%, para todos los modelos se rechaza la hipótesis de independencia y, la unión de ambos, la cobertura condicional resulta incorrecta para los seis modelos. Para el índice DAX, observamos que si se distribuyen los residuos según una exponencial generalizada asimétrica, los modelos ARMA y GARCH aceptan la hipótesis de cobertura incondicional correcta, aceptamos independencia para el modelo ARMA (según una exponencial generalizada asimétrica) y para el modelo APARCH (según una normal), en este caso sólo al 1%, la cobertura condicional correcta sólo se da en el caso del modelo ARMA (según una exponencial generalizada asimétrica). El índice DOWJONES, para los tres modelos se acepta la cobertura incondicional correcta sólo si se distribuyen según una exponencial generalizada asimétrica, aceptamos

independencia para el modelo APARCH y para el modelo GARCH, pero en el caso de estar distribuido según una exponencial generalizada asimétrica sólo aceptamos al 1%; y rechazamos que la cobertura condicional sea correcta salvo para el modelo GARCH (según una exponencial generalizada asimétrica), que se acepta al 1% y para el modelo APARCH (según una exponencial generalizada asimétrica), que se acepta tanto al 1% como al 5%. Para el EUROSTOXX, aceptamos cobertura incondicional correcta en los modelos ARMA (según una t de Student) y el modelo GARCH (según una t de Student asimétrica), la independencia es aceptada al 1% en los modelos ARMA (según una t de Student) y APARCH (según uno normal) y, la cobertura condicional es incorrecta en todos los modelos salvo en el ARMA (según una t de Student). El último índice, el NIKKEI, aceptamos cobertura incondicional correcta en el modelo ARMA (según una t de Student), aceptamos dependencia en los seis modelos y, la cobertura condicional es incorrecta también para los seis modelos.

Si analizamos las pérdidas de Gneiting y Raftery se observa que, en general, los modelos heteroscedásticos son los que mayor nivel de exactitud tienen debido a su mayor poder adaptativo a la evolución de la serie de rentabilidades (ver gráficos 14), siendo los modelos APARCH con errores normales (salvo el caso del índice BEL20) los que tienen un comportamiento más exacto, con la única excepción del índice DAX en el que el mejor comportamiento corresponde a un GARCH con errores normales.

6. Conclusiones

En el presente trabajo se ha estimado la volatilidad presente en la evolución diaria de 6 índices bursátiles mediante el uso de modelos ARMA-GARCH y ARMA-APARCH, analizando el comportamiento predictivo de los modelos ajustados, tanto en términos puntuales como de intervalos, así como el impacto ejercido no sólo por la volatilidad estimada sino también por la existencia del llamado efecto asimétrico y la falta de normalidad de las distribuciones condicionales del error.

Se observa, con carácter general, la existencia de una volatilidad cambiante en el tiempo, con altos niveles de persistencia, efectos asimétricos significativos (con un mayor impacto sobre la volatilidad de la sobre-predicción de la rentabilidad) así como falta de normalidad por la presencia de una fuerte leptocurtosis y una ligera asimetría negativa significativas. Todo ello provoca que, en general, los modelos con mejores niveles de ajuste a los datos son los APARCH con errores no normales.

A nivel predictivo puntual todos los modelos predicen de forma similar debido a la ausencia de dependencias en media significativas (todos los modelos ajustados son ruido blanco). Las diferencias se muestran en la predicción por intervalos en la que se observa que los modelos heteroscedásticos muestran un comportamiento más adaptativo a las oscilaciones de la serie dando lugar a intervalos más acurados tanto en términos de longitud como de menor nivel de error de infra y sobrepredicción. La incorporación del efecto asimétrico al modelo mejora ligeramente los niveles de exactitud no observándose un comportamiento muy diferenciado entre modelos con errores no normales y no normales. Estos comportamientos son muy interesantes de cara a estimar los llamados valores de riesgo (VaR) que se construyen a partir de los extremos inferior (en operaciones a largo) y superior (en operaciones a corto) de los intervalos de predicción de las rentabilidades lo cual justifica el éxito de este tipo de modelos en la evaluación diaria de los niveles de riesgo de las series financieras.

7. Bibliografía

- BOLLERSLEV, T. (1986) Generalized autorregressive conditional heteroscedasticity. *Journal of Econometrics*; 31: 307-327
- CHRISTOFFERSEN, P.F. (1998) Evaluating interval forecasts. *International Economic Review*, 4: 841-862
- DING, Z.; GRANGER, C.W.J. y ENGLE, R.F. (1993) A long memory property of stocks returns and a new model. *Journal of Finance*; 1: 83-106
- ENGLE, R.F. (1982) Autorregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*; 50: 987-1007
- ENGLE, R.F. y GONZÁLEZ-RIVERA, G. (1991) Semiparametric ARCH models. *The Journal of Business and Economic Statistics*; 9: 345-359
- ENGLE, R.E. y Ng, V. (1993) Measuring and testing the impact of news on volatility. *Journal of Econometrics*; 48: 1749-1778
- GLOSTEN, L.R.; JAGANNATHAN, R. y RUNKLE, D.E. (1993) On the relation between the expected value and the volatility of nominal excess return on stocks. *Journal of Finance*; 48: 1779-1801

GNEITING, T. y RAFTERY, A.E. (2007) Strictly proper scoring rules, prediction and estimation. *Journal of the American Statistical Association*, 477: 359-378

NELSON, D.B. (1991) Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*; 59: 347-370

SCHWARZ, G. (1978) Estimating the dimension of a model. *The Annals of Statistics*, 6: 461-464

TAYLOR, S.J. (1994) Modeling stochastic volatility: a review and comparative study. *Mathematical Finance*; 4: 183-204

TSAY, R.S. (2013) *An introduction to analysis of financial data with R*.First Edition. John Wiley & sons, Inc.