Span Econ Rev (2008) 10:169–196
DOI 10.1007/s10108-007-9030-6

REGULAR ARTICLE

Volatility and VaR forecasting in the Madrid Stock Exchange

Trino-Manuel Ñíguez

Published online: 21 June 2007
© Springer-Verlag 2007

Abstract This paper provides an empirical study to assess the forecasting performance of a wide range of models for predicting volatility and VaR in the Madrid Stock Exchange. The models performance was measured by using different loss functions and criteria. The results show that FIAPARCH processes capture and forecast more accurately the dynamics of IBEX-35 returns volatility. It is also observed that assuming a heavy-tailed distribution does not improve models ability for predicting volatility. However, when the aim is forecasting VaR, we find evidence of that the Student's *t* FIAPARCH outperforms the models it nests the lower the target quantile.

Keywords FIAPARCH · Heavy-tailed distributions · Leverage effect · Long memory · VaR

JEL Classification C32 · C52 · C53 · G15

1 Introduction

Parametric models for financial asset returns volatility have undergone great development since the seminal autoregressive conditional heteroscedasticity (ARCH), and generalized ARCH (GARCH) models of Engle (1982) and Bollerslev (1986). The literature on this topic is very extensive; see, for instance, Bollerslev et al. (1992) and Bauwens et al. (2006) for comprehensive surveys on univariate and multivariate ARCH-type models.

T.-M. Ñíguez (✉)
Department of Economics and Quantitative Methods, Westminster Business School,
University of Westminster, 35 Marylebone Road, London NW1 5LS, UK
e-mail: T.M.Niguez@wmin.ac.uk

In most of empirical work related to the volatility modeling of high-frequency financial data, it is found that the sum of the GARCH coefficients remains very near to the stationarity bound of one, which is interpreted as long-run dependence in volatility. To model this dynamics, integrated GARCH (IGARCH) processes were proposed by Engle and Bollerslev (1986) assuming directly non-stationarity. These processes present a constant cumulative impulse response function implying that shocks to the conditional variance persist indefinitely, which is not realistic; see, for instance, Bollerslev and Engle (1993). Recent contributions provide evidence on spurious integration in the conditional variance as an approximation to the true long-memory fractionally integrated behavior; see, for instance, Andersen and Bollerslev (1998) and Lobato and Savin (1998). On the other hand, we know that covariance-stationary GARCH processes generate autocorrelations that decrease excessively quickly in relation to the observed ones of the usual volatility proxies (absolute value or squared returns); see, for instance, Ding and Granger (1996).

In turns, fractionally integrated GARCH (FIGARCH) processes (Baillie et al. 1996), as other classes of long-memory models for the conditional variance, namely, long-memory stochastic volatility (Breidt et al. 1998) or long-memory GARCH (Granger and Ding 1996), are shown to describe well the observed slow decay in the autocorrelations of squared or absolute asset returns, usually interpreted as long-memory in financial market volatility.[1] Specifically, two are the main reasons why long-memory processes for the conditional variance have attracted a renewed interest: (a) the fact that it is certainly possible that from a pragmatic perspective, the assumption of long-memory may yield the most accurate empirical out-of-sample forecasts, (Bollerslev and Mikkelsen 1999), even if the data generating process shows structural breaks or week dependence; see, for instance, Diebold and Inoue (1999) and Morana and Beltratti (2004), and (b) volatility modeling and forecasting are the cornerstones for a broad variety of financial applications, such as: hedging strategy, option pricing, asset allocation and risk measurement and regulation, more specifically, e.g., for Value-at-Risk (VaR) forecasting; see, for instance, Jorion (2001) for further details, and Muller et al. (1997) for economic interpretations of long-run dependence in assets returns volatility. We also find in the literature some skeptical views about long-run dependence in volatility, which are based on that it is a spurious feature due to unaccounted structural changes; see, for instance, Ryden et al. (1998). On the other hand, more recently, Morana and Beltratti (2004) have shown that neglecting the change in volatility regimes is not relevant for short term forecasting once the process embodies a long-memory structure.

Turning to the volatility forecasting literature, the out-of-sample performance of short-memory GARCH-type of models has extensively been analyzed, for instance, Lee (1991) suggested that GARCH models that explain asymmetric responses of volatility to positive and negative shocks [leverage effect, (Black 1976)] provide more accurate forecasts than GARCH or IGARCH models. More recently, López and Walter

[1] It is worth noting that, although FIGARCH processes are an adaptation for the conditional variance of the autoregressive fractionally integrated moving average (ARFIMA) class of models (see, for instance, Hosking 1981), their theoretical properties differ significantly from those of their ARFIMA counterparts [see, for instance, Baillie (1996) for further details].

(2001) and Ferreira and López (2005) study the predicting ability of multivariate GARCH models under Gaussian and heavy-tailed distributions for volatility and VaR, suggesting that simple models seem to be more appropriate for VaR forecasting, in line with the results in Lucas (2000). On the other hand, Vilasuso (2002) and Ñíguez and Rubia (2006) find evidence in favor of Gaussian long-memory models for forecasting univariate and multivariate conditional variance. All of these empirical studies consider exchange rates and/or interest rates data.

Many are also the contributions in relation to stock-returns volatility modeling and forecasting; see, for instance, Brooks and Persand (2003) and Hansen and Lunde (2005) for contributions on volatility and VaR forecasting analysis of short-memory GARCH models, and Hansen and Lunde (2006) for a recent paper on volatility forecasting including Gaussian FIGARCH models and different measures of "realized" volatility. In particular, and focusing on research about the Madrid Stock Exchange (MSE), León and Mora (1999) studied the performance of a wide variety of short-memory volatility processes concluding that parametric models provide a better in-sample goodness-of-fit. On the other hand, Mármol and Reboredo (2000) and Arteche (2004) find evidence on the existence of long-memory in the IBEX-35 volatility by using nonparametric tests by Perron and Ng (1996) and Lobato and Robinson (1997), and local Whittle estimation of long-memory stochastic volatility models. However, up to the knowledge of the author much less research has been done about the usefulness of FIGARCH-type models including heavy-tailed distributions for forecasting stock-returns volatility and VaR, or for forecasting returns volatility and VaR, in particular, in the MSE market.

Whether fractional integration to account for long memory in volatility together with volatility powers, leverage effects, and heavy-tailed distributions helps to forecast stock-returns risk, is the central question this paper attempts to answer. It is worth noting that in the prior literature one or more of these features were not considered thus leaving the models comparison analysis incomplete. Thus, for the purpose of this study, an empirical analysis is performed to compare the forecasting ability for volatility and VaR of a broad range of short- and long-memory GARCH processes, including: GARCH, asymmetric GARCH (AGARCH) (Glosten et al. 1993), asymmetric power ARCH (APARCH) (Ding et al. 1993), exponentially weighted moving average (EWMA), FIGARCH and FIAPARCH (Tse 1998). On the other hand, it is widely known that (FI)GARCH-type processes are not enough to fully explain the leptokurtosis observed in asset returns distributions. So a heavy-tailed distribution should be assumed to account for this fact. We find evidence in the literature that shows that GARCH models with conditionally t-distributed innovations are able to capture all the observed excess kurtosis; see, for instance, the seminal paper of Bollerslev (1987). The empirical analysis, performed for the MSE index IBEX-35, considers thereof two distributions for the returns: the Gaussian (as a benchmark) and the Student's t.

It is worth mentioning that, the decision on the adequacy of the VaR methods for financial risk-modeling and risk-management and, on the most appropriate model for VaR forecasting are burning issues of current economic and econometric debate in academic and practitioners forums; see, for instance, Jorion (2002), Danielsson et al. (2004) and Ferreira and López (2005). In our empirical application VaR forecasts from the different models are validated by using several measures of performance, including: the regulatory loss functions proposed by López (1999), the HITs test of Engle and Manganelli (2004), and predictive quantile loss functions based on the methodology in Koenker and Bassett (1978).

The remainder of the paper is organized as follows. Section 2 presents the data. Section 3 provides an overview of the theoretical models for conditional heteroscedasticity. Section 4 discusses the estimation procedure and the models in-sample goodness-of-fit. Section 5 presents the results of the out-of-sample volatility forecasting. Section 6 sets up the framework and presents the results of the models evaluation for VaR forecasting. Finally, Sect. 7 gathers a summary of the conclusions and suggests topics for future research.

2 The data: stylized features of IBEX-35 returns

The data consists of daily closing prices of the MSE index IBEX-35, P_t, from January 07, 1987 to April 26, 2002, for a total of 3,993 observations. The series of continuously compounded daily returns on the index is given by, $r_t = \ln(P_t/P_{t-1})$. Figure 1 below depicts P_t and r_t.

Table 1 below summarizes the descriptive statistics of r_t. The measures of kurtosis and skewness indicate that the unconditional returns distribution is leptokurtic and slightly negatively skewed. The Jarque–Bera normality test null hypothesis is rejected at any reasonable significance level.

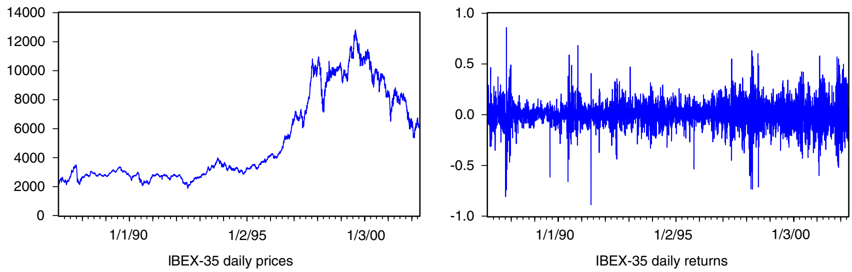

Fig. 1 Plots of IBEX-35 daily closing prices and returns

Table 1 Descriptive statistics for daily IBEX-35 returns, 7 January 1987–26 April 2002

Statistic	r_t
Sample size	3993
Mean	0.002527
Median	0.004819
Maximum	0.859183
Minimum	−0.887579
SD	0.135309
Skewness	−0.273432
Kurtosis	7.142843
Jarque–Bera	2905.274
P value	0.000000

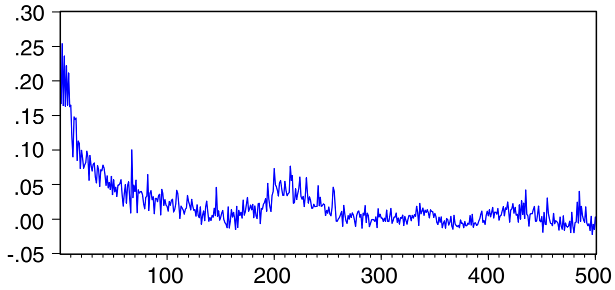

Fig. 2 Plot of autocorrelations of IBEX-35 daily squared returns, r_t^2 (500 lags)

Turning to the r_t series, a small predictable component is detected in the level, which is filtered by using a moving average process of order one, MA(1), selected according to the Box and Jenkins' methodology and the Akaike Information Criterion (AIC hereafter).[2]

In Fig. 1, we observe the usual clusters in returns volatility (Mandelbrot 1963) meaning that asset returns are not independent. The Lagrange–Multiplier test for ARCH effects (Engle 1982) rejects the null at any reasonable significance level even for high long lags, which is interpreted as long-run temporal dependence in volatility. This fact can also be noted graphically by observing the hyperbolic decay of the r_t^2 autocorrelations, depicted in Fig. 2.

Another important stylized fact of financial assets returns that any valuable model should convincingly explain, is the so-called leverage effect: "A fall of a firm's stock value relative to the market value of its debt causes a rise in its debt-equity ratio and increases its stock volatility" (Black 1976). As a consequence, volatility tends to increase less in response to 'good news' (excess returns higher than expected) than in response to 'bad news' (excess returns lower than expected). GARCH models can be generalized to account for this effect, as explained in the next section.

3 Modeling conditional volatility

This section aims to provide an overview of the development undergone by univariate conditional heteroscedasticity models designed for modeling financial asset returns volatility.

Define formally the univariate process for the stock returns, r_t, as follows,

$$r_t = \mu_t + \varepsilon_t, \quad \varepsilon_t = \sigma_t \eta_t \tag{1}$$

$$\mu_t = \mu(\Omega_{t-1}; \theta), \tag{2}$$

$$\sigma_t^2 = \sigma^2(\Omega_{t-1}; \theta), \tag{3}$$

[2] The existence of small linear dependences in the level of index returns has been attributed to non-synchronous trading in the stocks that compose the index, in accordance with the efficient market hypothesis; see, for instance, Sentana and Wadhwani (1992) for a comprehensive study on this topic.

where $\mu(\cdot)$ and $\sigma^2(\cdot)$ are functions known up to a finite-dimensional vector of true parameter values $\theta \in \mathbb{R}^P$, Ω_{t-1} is the information set available at time $t-1$, ε_t is the innovation, σ_t is the latent (unobservable) volatility, and η_t is a martingale difference sequence satisfying $E(\eta_t|\Omega_{t-1};\theta) = 0$ and $V(\eta_t|\Omega_{t-1};\theta) = 1$. As a consequence,

$$\mathbf{E}(r_t|\Omega_{t-1};\theta) = \mu_t, \tag{4}$$

$$\mathbf{V}(r_t|\Omega_{t-1};\theta) = \sigma_t^2. \tag{5}$$

The ARCH of order p (ARCH(p)) model of Engle (1982) was proposed to explain clustering dynamics in ε_t^2. The model is given by the following equation,

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2, \tag{6}$$

where $\omega > 0$, $\alpha(L) = \sum_{k=1}^{p} L^k \varepsilon_{t-k}^2$ with $\alpha_k \geq 0\ \forall k$.

The GARCH model was proposed by Bollerslev (1986) to avoid the high-order ARCH specification demanded to capture the conditional variance behavior. The GARCH of indexes p, q (GARCH(p, q)) model is expressed as,

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2, \tag{7}$$

where $\beta(L) \equiv \sum_{s=1}^{q} \beta_s L^s$ with $\beta_s \geq 0\ \forall s$. All the roots of $\alpha(L)$ and $[1-\alpha(L)-\beta(L)]$ have to be outside the unit circle to ensure the covariance stationarity of the process. This model can also be expressed in ARMA(m, p) form for ε_t^2 as follows,

$$[1-\alpha(L)-\beta(L)]\varepsilon_t^2 = \omega + [1-\beta(L)]\upsilon_t, \tag{8}$$

where $m = \max\{p, q\}$, and $\upsilon_t = \varepsilon_t^2 - \sigma_t^2$ are martingale differences and usually are interpreted as the innovations to the conditional variance (cf., e.g., Robinson 2003). When the polynomial $[1-\alpha(L)-\beta(L)]$ contains a unitary root we have the IGARCH(p, q) process (Engle and Bollerslev 1986), which in its ARMA($m-1$, p) form is given by,

$$\phi(L)(1-L)\varepsilon_t^2 = \omega + [1-\beta(L)]\upsilon_t, \tag{9}$$

where $\phi(L) \equiv [1-\alpha(L)-\beta(L)](1-L)^{-1}$ is of order $m-1$.

Baillie et al. (1996) proposed the FIGARCH of orders p, d, q (FIGARCH(p, d, q)) model to capture long-memory in asset returns volatility. The model is obtained by replacing the differencing operator in Eq. (9) with the fractional polynomial $(1-L)^d$. The process in its ARFIMA form of orders $m-1, d, q$ (ARFIMA($m-1$,d,q)) for ε_t^2 is defined as,

$$\phi(L)(1-L)^d\varepsilon_t^2 = \omega + [1-\beta(L)]\upsilon_t, \tag{10}$$

where $0 \leq d \leq 1$. An alternative representation for the FIGARCH(p, d, q) model is the following,

$$\sigma_t^2 = \omega[1-\beta(1)]^{-1} + \left[1-\phi(L)[1-\beta(L)]^{-1}(1-L)^d\right]\varepsilon_t^2, \tag{11}$$

with cumulative impulse response weights (CIRW) given by the coefficients in the lag polynomial, $\lambda(L)$,

$$\lambda(L) = 1 - [(1 - \beta(L)]^{-1}\phi(L)(1 - L)^d. \tag{12}$$

The fractional differencing operator, $(1-L)^d$, has a binomial expansion which is most conveniently expressed in terms of the hypergeometric function as follows,

$$(1 - L)^d = F(-d, 1; 1; L) = \sum_{k=0}^{\infty} \frac{\Gamma(k - d)}{\Gamma(k + 1)\Gamma(-d)} L^k. \tag{13}$$

This process, as the IGARCH, is not covariance-stationary but it is strictly stationary and ergodic for $0 \leq d \leq 1$, as argued in Bollerslev and Mikkelsen (1996) using a dominance-type argument, although more recently Giraitis et al. (2000) have suggested that further research on this direction is needed. FIGARCH models CIRW decay hyperbolically towards zero capturing the very slow decay observed in the autocorrelations of the returns volatility proxies. From a forecasting perspective this feature is interpreted as: shocks affect the optimal forecast of the FIGARCH conditional variance over long lags, unlike short-memory GARCH processes for which the influence of shocks disappears exponentially or, IGARCH processes in which shocks show infinite persistence.

On the other hand, Ding et al. (1993) generalized the GARCH process in Eq. (7), by considering the modeling of a power of the conditional standard deviation, which is a function of past residuals (to account for the leverage effect) and powers of past conditional standard deviations. The process is called APARCH(p, q) and is defined as,

$$\sigma_t^\delta = \omega + \sum_{k=1}^{p} \alpha_k \left(|\varepsilon_{t-k}| - \gamma_k \varepsilon_{t-k}\right)^\delta + \sum_{s=1}^{q} \beta_s \sigma_{t-s}^\delta, \tag{14}$$

where $\delta > 0$ and $|\gamma| < 1$. The term $(|\varepsilon_t| - \gamma\varepsilon_t)^\delta$ accounts for the asymmetric response of volatility to positive and negative shocks when $\gamma \neq 0$. Thus, when $0 < \gamma < 1$ a positive innovation increases volatility less than a negative one, and vice versa for $-1 < \gamma < 0$. When $\gamma = 0$ a positive innovation has the same effect on volatility as a negative one of the same magnitude. This model nests the AGARCH(p, q) of Glosten et al. (1993) when $\delta = 2$ (see Appendix A for further details). See, for instance, Ling and McAleer (2002) for a comprehensive analysis on the statistical properties of APARCH processes.

The FIAPARCH(p, d, q) model combines the FIGARCH(p, d, q) and the APARCH (p, q) and can be defined as follows,

$$\sigma_t^\delta = \frac{\omega}{1 - \beta(1)} + \left[1 - \phi\,(L)\,[1 - \beta\,(L)]^{-1}\,(1 - L)^d\right](|\varepsilon_t| - \gamma\varepsilon_t)^\delta. \tag{15}$$

This process is not covariance-stationary, as the FIGARCH(p, d, q) and IGARCH (p, q) processes. Note that since the hypergeometric function $F(-d, 1, 1; L)$ evaluated at $L = 1$ equals zero we obtain that $\lambda(1) = 1$ so the process present a unit root (see Gradshteyn and Ryzhnik (1980) for a complete survey on hypergeometric functions). See also Appendix A for a review of the nested models in the FIAPARCH structure.

It is worth mentioning that, an alternative process to the FIAPARCH would be the exponential FIGARCH (FIEGARCH) of Bollerslev and Mikkelsen (1996). This process is specified in terms of the logarithm of the conditional variance, and as the FIAPARCH, accounts for the leverage effect in stock returns. Furthermore, it allows for strict covariance-stationarity and ergodicity, and does not require parameter constraints for the model to be well-defined. In this study, we choose the FIAPARCH instead of the FIEGARCH since we are interested in exploring the forecasting performance of an asymmetric FIGARCH process that furthermore presents the flexibility of leaving the power of the conditional standard deviation free to be estimated from the data.

In this paper the first-order version of the aforementioned processes is considered: FIAPARCH$(1, d, 1)$, FIGARCH$(1, d, 1)$, APARCH$(1, 1)$, AGARCH$(1, 1)$, GARCH $(1, 1)$ and EWMA$(1, 1)$. In particular, it is worth noting that the EWMA model implemented in the RiskMetrics program (Morgan 1996) for VaR forecasting corresponds to a particular case of an IGARCH$(1, 1)$ model without drift and with fixed coefficient β between 0.94 and 0.97 . In the EWMA model in this paper, the weight, β, is a parameter to be estimated. The model is given by,[3]

$$\sigma_t^2 = (1-\beta){\varepsilon_{t-1}}^2 + \beta\sigma_{t-1}^2. \tag{16}$$

The CIRW sequence, $\{\lambda_k\}_{k=0}^{\infty}$, for the FIGARCH$(1, d, 1)$-type of processes can be expressed as follows,

$$\lambda_k = \begin{cases} 0, & k=0 \\ d-\beta+\phi, & k=1 \\ \beta\lambda_{k-1} + \left(\frac{k-1-d}{k} - \phi\right)\delta_{k-1}, & k \geq 2, \end{cases} \tag{17}$$

with

$$\delta_k = d\frac{\Gamma(k-d)}{\Gamma(1-d)\,\Gamma(k+1)} \quad \forall k \geq 1. \tag{18}$$

For the FIGARCH-type of processes to be well-defined and the conditional variance positive almost surely for all t, all CIRW must be non-negative. In this paper we use the sufficient conditions provided by Bollerslev and Mikkelsen (1996) for the FIGARCH$(1, d, 1)$ specification,

$$-d \leq \phi \leq \frac{2-d}{3}, \quad d\left(\phi - \frac{1-d}{2}\right) \leq \beta\big(\phi-\beta+d\big), \tag{19}$$

It is worth mentioning that other sets of sufficient parameter constraints for this model have also been proposed by Baillie (1996) and Chung (1999). In the former case, those conditions are stronger, while in the latter they are complementary; i.e., weaker only

[3] This model provides a reasonable approximation of an IGARCH model because in empirical estimation for asset returns, the estimate of ω is, although significant, very near zero.

for a subset of the parameter space. More recently, Conrad and Haag (2006) have proposed necessary and sufficient conditions which enlarge the feasible parameter set in relation to the aforementioned ones, thus likely providing the model with greater flexibility to capture and forecast the dynamics in the conditional variance. The study of the forecasting performance of FIGARCH models under Conrad and Haag (2006) conditions constitutes an interesting topic for future research.

4 Estimation and in-sample analysis

White (1982) and Weiss (1986) showed that consistent and asymptotically normal parameter estimates of conditional mean and finite ARCH processes can be obtained by means of Maximum Likelihood (ML) estimation under normality, providing that the unconditional mean and variance are well-defined. These estimates obtained under possible distribution misspecification (i.e., when the data are, in fact, not normally distributed) are called Quasi-Maximum Likelihood estimates (QMLE), which keep the aforementioned properties but they are not necessarily asymptotically efficient (Bollerslev and Wooldridge 1992). More recently, Straumann and Mikosch (2006) have proven the consistency and asymptotic normality of the QMLE for AGARCH models. Unfortunately, consistency of MLE when another distribution rather than the Normal is assumed is not guaranteed; see Newey and Steigerwald (1997) for further details. On the other hand, it is worth noting that since for the FIGARCH$(1, d, 1)$ model the population variance does not exist, consistency and asymptotic normality is not guaranteed for the (Q)MLE. Nonetheless, Baillie et al. (1996) argued, through a dominance-type argument, that for the FIGARCH$(1, d, 0)$ model the QMLE is consistent and asymptotically normal since it is for the IGARCH$(1, 1)$ process. They also provided evidence on the good performance of the QMLE for FIGARCH$(1, d, 1)$ processes by means of simulation techniques; see also Caporin (2003). Hence, standard asymptotic inference regarding (Q)MLE of FIGARCH processes can be carefully used in practice. It is also worth noting that asymptotic properties of (Q)MLE for APARCH and FIAPARCH processes have not been formally established yet, remaining this topic as an interesting area for further research.[4]

Being aware of the limitations in relation to the estimation procedures, the aforementioned conditionally heteroscedastic models are compared according to their in- and out-of-sample forecasting performance. The estimation is carried out following the two-stages procedure discussed in White (1994), thus the sample is split into two parts: (a) the in-sample period, which takes the first $T = 3,093$ observations and, (b) the out-of-sample period, which includes the last $N = 900$ observations.

Stage 1. The residuals series, $\widehat{\varepsilon}_t$, is obtained by estimating a MA(1) process by Ordinary Least Squares.

Stage 2. The $\widehat{\varepsilon}_t$ series is then used to estimate the parameters of the models. So, QMLE are obtained by maximizing the sample log likelihood corresponding to the Gaussian distribution, and MLE when the standardized Student's t distribution is assumed, which is given by,

[4] See Hidalgo (1997) for nonparametric estimation techniques under long-run temporal dependence.

$$\mathcal{L}_t(\Upsilon/\Omega_{t-1}) = T\ \ln\left[\frac{\Gamma[(\xi+1)/2\xi]}{\pi^{1/2}\Gamma(1/2\xi)\left(\frac{1-2\xi}{\xi}\right)^{1/2}}\right] \\ -\sum_{t=1}^{T}\frac{\ln(\sigma_t^2)}{2}-\frac{\xi+1}{2\xi}\sum_{t=1}^{T}\ln\left[1+\frac{\xi\varepsilon_t^2}{(1-2\xi)\sigma_t^2}\right] \tag{20}$$

where Υ is a finite-dimension unknown parameter vector, and ξ is the inverse of the degrees of freedom parameter.[5]

To apply the procedure to FIGARCH-type models, the infinite lag polynomial was truncated at the usual lag of 1, 000. All ML estimations were carried out under the non-negativity parameter constraints discussed in Sect. 3 using the cml library of GAUSS. Robust standard errors in the second stage were computed following Bollerslev and Wooldridge (1992). Furthermore, we carried out a further Newton–Raphson iteration from consistent estimators to obtain estimators asymptotically equivalent to joint QMLE; see Pagan (1986) for further details on this procedure.

Table 2 reports the estimation results.[6] Consistent with the prior literature, the estimates for $\widehat{\alpha}+\widehat{\beta}$ are close to unity, which indicates high persistence in the IBEX-35 returns volatility. The degrees of freedom coefficient, $\widehat{\nu}$, is around 7, even after correcting for volatility clustering, confirming the existence of leptokurtosis in the returns conditional distribution. The asymmetric parameter estimate, $\hat{\gamma}$, is statistically different from zero or one in all asymmetric models, confirming the existence of asymmetric responses to positive and negative shocks in the index return volatility. The power parameter estimate, $\hat{\delta}$, is statistically different from one or two in APARCH-t and FIAPARCH models, while for the APARCH-n, $\hat{\delta}$ is not statistically different from 2 suggesting that the AGARCH-n specification is more appropriate than the APARCH-n given its simpler parameterization. The long-memory parameter estimate, $\widehat{d} \in [0.576, 0.654]$, is statistically different from zero or one for all models, which confirms the existence of a strong long-run persistence (long-memory) in the volatility of the IBEX-35 returns, in accordance with prior results reported by Mármol and Reboredo (2000) and Arteche (2004).

The AIC and the likelihood ratio (LR) test are used to evaluate the in-sample goodness-of-fit of the models.[7] According to the AIC, all models under the Student's t distribution provide a better fit to the data. Furthermore, the greatest AIC from the Student's t models ($\mathrm{AIC}_{\mathrm{EWMA}\text{-}t} = -1.6971$) is still smaller than the one of the most flexible Gaussian model, i.e., the FIAPARCH-n, ($\mathrm{AIC}_{\mathrm{FIAPARCH}\text{-}n} = -1.6206$). Therefore, we can conclude that the FIAPARCH-t is the model that presents the most accurate in-sample fit. The same conclusion is obtained by performing the LR test,

[5] Note that, since the Student's t distribution approaches the Normal as the degrees of freedom parameter tends to infinite, $\nu \to \infty$, it is more convenient to use its inverse, i.e., $\xi = 1/\nu$, as a measure of the tail thickness, which will always remain in the finite range of $0 \le \xi < \frac{1}{2}$; see, for instance, Fiorentini et al. (2003).

[6] For the sake of simplifying notation, hereafter 'n' and 't' preceded by the conditional variance process denote that the assumed distribution for the returns is either the Normal or the Student's t, respectively.

[7] The AIC statistic is defined as $2(\varrho - \log L)/(T-N)$, where ϱ is the number of estimated parameters. Note that a smaller AIC corresponds to a better fit.

Table 2 Estimation results

	GARCH	AGARCH	APARCH	EWMA	FIGARCH	FIAPARCH
Gaussian						
ω	0.000 (2.55)	0.000 (2.54)	0.000 (1.73)		0.001 (1.43)	0.015 (1.86)
α	0.132 (5.95)	0.127 (5.37)	0.128 (5.56)	0.066 (3.69)		
β	0.824 (29.7)	0.824 (28.0)	0.827 (23.0)	0.934 (51.8)	0.171 (1.63)	0.232 (2.83)
ϕ					0.636 (2.05)	0.698 (10.7)
γ		0.148 (2.03)	0.154 (1.94)			0.292 (2.38)
δ			1.914 (5.65)			1.242 (7.46)
d					0.584 (1.45)	0.576 (5.45)
log L	2485.7	2493.0	2493.1	2327.8	2485.7	2512.1
AIC	−1.6052	−1.6094	−1.6089	−1.5046	−1.6047	−1.6206
Student's t						
ω	0.000 (2.14)	0.000 (2.94)	0.001 (2.46)		0.000 (2.59)	0.006 (1.61)
α	0.142 (7.00)	0.140 (6.94)	0.143 (8.08)	0.091 (6.94)		
β	0.849 (41.5)	0.849 (40.9)	0.865 (46.5)	0.909 (68.6)	0.142 (2.73)	0.153 (3.48)
ϕ					0.634 (8.50)	0.681 (9.01)
γ		0.078 (2.15)	0.098 (2.25)			0.140 (2.87)
δ			1.402 (9.28)			1.465 (7.30)
d					0.628 (8.05)	0.654 (7.90)
ν	6.791 (6.93)	6.834 (6.91)	6.732 (6.99)	7.404 (8.43)	7.018 (9.09)	7.505 (6.58)
log L	2653.8	2655.8	2661.1	2626.6	2661.9	2669.8
AIC	−1.7134	−1.7141	−1.7168	−1.6971	−1.7180	−1.7218

The reported coefficients shown in each row of the table are (Q)MLE of Gaussian and Student's t GARCH-(1,1)-type and FIGARCH(1,d,1)-type models for MA(1) innovations from IBEX-35 daily returns. Robust (Q)ML t -statistics are reported in parenthesis next to the parameter estimates. AIC denotes the Akaike Information Criterion and log L the value of the log-likelihood function

since the $\text{LR}_{\gamma=0,\delta=2}$ test statistic equals 76 and the null is rejected at any reasonable significance level of the corresponding asymptotic χ^2 distribution.

It is also remarkable that the modeling of the power of the conditional standard deviation provides the models with a greater ability to fit the data only under t-distributed errors, since the APARCH-t model adjusts better than the AGARCH-t process, but curiously, the opposite occurs under Gaussianity. Furthermore, GARCH-n, AGARCH-n and APARCH-n models perform slightly better than the FIGARCH-n. Finally, it is worth noting that the FIGARCH-t provides a much better fit than the short-memory GARCH-t models, as expected.

In summary, we find that the t-distribution enhances the in-sample forecasting capacity of the models, especially when models are able to capture the long-memory feature of the data. As a conclusion, the FIAPARCH -t and the EWMA-n models show the best and the worst in-sample goodness-of-fit, respectively, as it is inferred from their AIC.

5 Out-of-sample volatility forecasting

The out-sample forecasts evaluation of the models is outlined as follows:

First, the models estimation is repeated N times by using a rolling window to generate N 1 day ahead volatility forecasts. At each iteration, the in-sample window is updated (i.e., a new observation enters the sample and the oldest one is removed) so that the window length remains constant. The 1 day ahead forecast error from model m, denoted as $e_{m,T+i}$, $(i = 1, \ldots, N)$ is given by

$$e_{m,T+i} = \widehat{\varepsilon}^2_{T+i} - \widehat{\sigma}^2_{m,T+i} \tag{21}$$

where $\widehat{\sigma}^2_{m,T+i}$ is the 1 day projection of the conditional variance from model m given the information at time $T + i - 1$, $\mathbf{E}_{T+i-1}(\varepsilon^2_{T+i}) = \widehat{\sigma}^2_{m,T+i}$, and $\widehat{\varepsilon}^2_{T+i}$ is the chosen proxy for the unobservable "true" volatility,[8] which is given by,

$$\widehat{\varepsilon}^2_{T+i} = \left[r_{T+i} - \mathbf{E}_{T+i-1}(r_{T+i})\right]^2, \tag{22}$$

$$\mathbf{E}_{T+i-1}(r_{T+i}) = \widehat{\mu} + \widehat{\theta}\widehat{\varepsilon}_{T+i-1}. \tag{23}$$

Second, several criteria are used to assess the forecasting performance of the models, including:

(i) The Mincer–Zarnowitz regression (Mincer and Zarnowitz 1969), which consists on estimating the following equation,

$$\widehat{\varepsilon}^2_{T+i} = \vartheta_0 + \vartheta_1\widehat{\sigma}^2_{m,T+i} + u_{T+i}. \tag{24}$$

Thus, the forecast from model m is optimal with respect to the available information set (Ω_{T+i-1}) if the null H_0: $(\vartheta_0, \vartheta_1) = (0, 1)$ is accepted.

(ii) Symmetric loss functions as: mean squared prediction error (MSPE), and mean absolute prediction error (MAPE); and the following asymmetric loss functions: mean mixed error of underprediction (MME(U)), and mean mixed error of overprediction (MME(O)), which penalize more heavily under- and over-predictions, respectively,

$$\text{MME(U)}_m = \frac{1}{N}\left(\sum_{i=1}^{N_U}\sqrt{\left|e_{m,T+i}\right|} + \sum_{i=1}^{N_O}\left|e_{m,T+i}\right|\right), \tag{25}$$

$$\text{MME(O)}_m = \frac{1}{N}\left(\sum_{i=1}^{N_U}\left|e_{m,T+i}\right| + \sum_{i=1}^{N_O}\sqrt{\left|e_{m,T+i}\right|}\right), \tag{26}$$

[8] Andersen and Bollerslev (1998) have shown that for high frequency asset returns a better proxy for the "true" volatility can be obtained as the sum of the squared intradaily returns. In our case, as in many empirical works, a problem arises concerning data availability.

Table 3 Mincer-Zarnowitz regression test

	Gaussian		Student's t	
	P value	R^2_m	P value	R^2_m
GARCH	0.072	0.1145	0.018	0.1156
APARCH	0.053	0.1322	0.142	0.1291
AGARCH	0.055	0.1383	0.244	0.1331
EWMA	0.073	0.1142	0.015	0.1195
FIGARCH	0.665	0.1101	0.032	0.1167
FIAPARCH	0.294	0.1447	0.763	0.1400

The table reports P values for the Mincer and Zarnowitz (1969) regression test, H_0: $(\vartheta_0, \vartheta_1) = (0, 1)$, and regression determination coefficients for each model m, denoted as R^2_m

where N_U is the number of underpredictions and N_O its complementary, and the probability of underprediction (PUnd), which can be estimated by N_U/N; see Brailsford and Faff (1996) for applications of these measures to asset returns volatility forecasts.

Finally, the Diebold and Mariano (1995) (DM henceforth) test is used to verify whether the difference between the loss functions from the different models is statistically significant. See Appendix B for further details on the implementation of this test in the context of this paper.

The results in Table 3 show that the null of optimal forecast is accepted in most of the cases for a significance level of 5%; P values are smaller than 0.05 only for GARCH-t and EWMA-t. But for higher significance levels, for instance, 10%, the null is accepted only for APARCH-t, AGARCH-t, FIGARCH-t and FIAPARCH models. In relation to the determination coefficients, it is observed that the percentage of explained out-of-sample volatility is small in all cases; between 11 per cent (GARCH, EWMA and FIGARCH) and 14% (FIAPARCH).[9] As a conclusion, the assumption of asymmetry seems to be more determinant than either heavy-tailed distribution or fractional integration ones. Furthermore, it is observed a better goodness-of-fit when the model embodies all former assumptions (i.e., in the FIAPARCH-t case).

Tables 4 and 5 present loss functions values for volatility forecasting and their corresponding DM test t-statistics, respectively. A sharp result that emerges from those tables is that the distributional assumption does not help to provide more accurate volatility forecasts. As regards the MSPE results, it is observed that asymmetric models provide systematically lower MSPE. Furthermore, FIAPARCH models yield the smallest significant MSPE, with no significant MSPE differences from FIAPARCH models under both distributions. On the other hand, as regards the MAPE, it is observed that GARCH-n, AGARCH-n and APARCH-n models yield a smaller MAPE than FIAPARCH processes, although differences are not statistically significant for the case of the GARCH-n.[10]

[9] It is worth mentioting that our R^2 are slightly higher than those found in the literature, which are usually lower than 10% (see, for instance, Andersen and Bollerslev 1998).

[10] The DM test t-statistics for the difference between the MAPE from FIAPARCH- t model and GARCH-n, AGARCH-n and APARCH-n models are −1.35, −4.96 and −5.66, respectively. Note that these statistics are not reported in Table 5, but the same conclusion can be extracted by using the statistics provided in that table for the models involved.

Table 4 Out-of-sample volatility forecasting loss functions

	MSPE	MAPE	MME(O)	MME(U)	PUnd
Gaussian					
GARCH	0.1768	0.2524	0.0920	0.0713	0.357
AGARCH	0.1733	0.2487	0.0908	0.0710	0.361
APARCH	0.1721	0.2485	0.0914	0.0705	0.353
EWMA	0.1768	0.2604	0.0985	0.0690	0.326
FIGARCH	0.1768	0.2587	0.0987	0.0687	0.321
FIAPARCH	0.1701	0.2534	0.0978	0.0675	0.318
Student's *t*					
GARCH	0.1771	0.2613	0.0976	0.0696	0.338
AGARCH	0.1736	0.2575	0.0966	0.0691	0.342
APARCH	0.1722	0.2570	0.0973	0.0686	0.331
EWMA	0.1764	0.2608	0.0978	0.0693	0.333
FIGARCH	0.1766	0.2628	0.0995	0.0689	0.326
FIAPARCH	0.1708	0.2547	0.0966	0.0685	0.335

The table reports loss functions for 1 day ahead conditional variance out-of-sample forecasts. The models forecasting ability is measured by the following loss functions: mean squared prediction error (MSPE), mean absolute prediction error (MAPE), mean mixed error of underprediction (MME(U)), mean mixed error of overprediction (MME(O)) and probability of underprediction (PUnd)

As for the asymmetric loss functions, the results show that differences are statistically significant except for some few cases, for instance, FIGARCH-*n* versus EWMA-*n* and GARCH-*t* versus EWMA-*t* . Firstly, we observed that Student's *t* models present a higher MME(O) and lower MME(U), so they tend to overpredict volatility. Secondly, focusing on the following four selected models because of their significantly smaller MSPE and MAPE: FIAPARCH models, AGARCH-*n* and APARCH-*n*, we observe that FIAPARCH models tend to overpredict volatility since they present a higher MME(O) and, a lower MME(U) and PUnd. Thirdly, EWMA-*n*, FIGARCH and FIAPARCH processes show the largest MME(O) and, the lowest MME(U)s and PUnds.

These findings can be summarized as follows: (a) the assumption of Student's *t* innovations does not contribute to get more accurate volatility forecast, (b) asymmetric models tend to show a better performance than symmetric ones. These results are generally consistent with those of Brooks and Persand (2003). (c) In particular, the EWMA, GARCH and FIGARCH models produce slightly but significantly the most erratic forecasts. Note that these results are in accordance with those obtained from the Mincer–Zarnowitz regression test. Furthermore, (d) FIAPARCH models tend to overpredict more and underpredict less volatility in relation to other accurate models, such as AGARCH-*n* and APARCH-*n*.

The results presented so far confirm the well-known fact of that a better in-sample fit does not guarantee more accurate predictions (see, for instance, Nelson 1976). However, from the perspective of practitioners in the financial markets, the out-of-sample analysis is of greater interest. For instance, risk managers and regulators must decide on the appropriate model for predicting the risk of investment portfolios

Table 5 DM test statistics for volatility forecasting

	MSPE	MAPE	MME(O)	MME(U)
Gaussian				
FIGARCH versus GARCH	−0.01*	6.63	17.83	−7.92
FIGARCH versus AGARCH	1.37	5.85	13.82	−4.74
FIGARCH versus APARCH	1.95	6.32	13.26	−3.95
FIGARCH versus EWMA	−0.02*	−0.95*	0.40*	−0.84*
FIGARCH versus FIAPARCH	2.72	3.13	1.47	2.35
GARCH versus AGARCH	1.50	2.76	2.65	0.80
GARCH versus APARCH	2.08	2.97	1.19*	2.30
GARCH versus EWMA	−0.02*	−3.81	−8.98	4.05
GARCH versus FIAPARCH	2.57	−0.55*	−8.76	6.64
FIAPARCH versus EWMA	−2.00	−2.75	−0.69*	−2.17
Student's *t*				
FIGARCH versus GARCH	−0.50*	2.12	7.03	−3.63
FIGARCH versus AGARCH	1.58	3.80	6.39	−0.68
FIGARCH versus APARCH	2.39	4.21	4.85	0.61
FIGARCH versus EWMA	0.17*	1.59	3.91	−1.17*
FIGARCH versus FIAPARCH	2.55	4.83	5.34	0.71*
GARCH versus AGARCH	1.73	3.12	2.68	1.49
GARCH versus APARCH	2.45	3.32	0.72*	2.92
GARCH versus EWMA	0.52*	0.42*	−0.41*	1.00*
GARCH versus FIAPARCH	2.40	3.72	1.69	2.35
FIAPARCH versus EWMA	−1.90	−2.83	−1.61	−1.23*
Cross comparison for FIAPARCH				
FIAPARCH-*n* FIAPARCH-*t*	−0.50*	−1.19*	3.32	−3.41

The table reports Diebold and Mariano test *t*-statistics for 1 day ahead conditional variance forecast. GARCH and the FIGARCH are taken as the reference models. The loss functions considered are the following: mean squared prediction error (MSPE), mean absolute prediction error (MAPE), mean mixed error of underprediction (MME(U)) and mean mixed error of overprediction (MME(O)). Statistics marked with an asterisk indicate that the null hypothesis of equal predictive ability is not rejected at 10% level

(usually represented by VaR measures), on which either updating decisions on capital allocation or setting up regulatory capital.

6 VaR analysis

VaR measures are key inputs in financial risk-management and risk-regulation. Their relevance arises from the Market Risk Amendment (1996) to the Basel Capital Accords of 1988. These agreements aim to reinforce international banking system stability by means of more strict external regulation (see, for instance, Jorion 2001 for a comprehensive overview of this subject). Thus, the Basel Committee for Banking Supervision (at the Bank for International Settlements) requires financial institutions with significant trading activities to report periodically the VaR of their investment portfolios, according to which the regulatory capital that the institution must hold to cover its exposure to market risk is established. Regulators evaluate banks VaR models by observing when the corresponding portfolio returns exceed the reported VaR, determining whether these VaR forecasts are 'acceptably accurate'. It is known that the evaluation methods currently used are questionable and can be improved. Those methods (binomial and interval forecast) are based on tests whose statistical power has been shown to be too low for discriminating among reasonable alternative models. VaR methodology for forecasting and regulating risk has attracted the attention of many researchers during these last years given its potential effects on competition in financial markets, efficient allocations of capital and economic general equilibrium paths (see, for instance, Danielsson et al. 2004). In relation to VaR models validation we find in the literature different techniques aiming to provide sufficient power for discriminating among similarly performing models. In this section, we apply the regulatory loss functions criteria proposed by López (1999), the predictive quantile loss function (see, for instance, Komunjer 2005) and the HIT or dynamic quantile test of Engle and Manganelli (2004).

6.1 VaR definition

VaR is defined as the maximum expected loss in the value of a portfolio, for a given probability α and a determined forecast horizon. The great popularity that this instrument has achieved among financial practitioners and regulators is essentially due to its conceptual simplicity; VaR reduces the market risk associated with any portfolio to just one number. We denote the cumulative assumed conditional distribution function of r_{T+i} for model m, as $F^m(r_{T+i}|\Omega_{T+i-1})$, then the $VaR^{\alpha}_{m,T+i}$ of a long position is defined as,

$$\alpha = \Pr\left[r_{T+i} \leq VaR^{\alpha}_{m,T+i}|\Omega_{T+i-1}\right] = F^m(VaR^{\alpha}_{m,T+i}|\Omega_{T+i-1}). \tag{27}$$

Equivalently, $VaR^{\alpha}_{m,T+i}$ can be defined as the solution to,

$$\int_{-\infty}^{VaR^{\alpha}_{m,T+i}} f^m(r_{T+i}|\Omega_{T+i-1})dr_{T+1} = \alpha, \tag{28}$$

where $f(r_{T+i}|\Omega_{T+i-1})$ is the assumed conditional density function of r_{T+i}. This definition states that the probability of a loss (r_{T+i}) greater or equal to the 1 day ahead VaR predicted at time $T+i-1$ from model m is α. In other words, $VaR^{\alpha}_{m,T+i}$ is the $(100 \cdot \alpha)$ per cent quantile of $f(r_{T+i}|\Omega_{T+i-1})$; $VaR^{\alpha}_{m,T+i} = f^{-1}(\alpha, r_{T+i}|\Omega_{T+i-1})$.

6.2 VaR forecasting framework

Taking 1 day as the forecast horizon, the VaR forecast for confidence level α, obtained at time $T+i-1$ from model m is given by,

$$\widehat{VaR}^{\alpha}_{m,T+i} = \widehat{\mu}_{T+i} - \widehat{z}^{\alpha}_{m,T+i-1} \cdot \widehat{\sigma}_{m,T+i}, \tag{29}$$

where $\widehat{\mu}_{T+i}$ and $\widehat{\sigma}_{m,T+i}$ are 1 day ahead forecasts of the conditional mean and conditional standard deviation, respectively, and $\widehat{z}^{\alpha}_{m,T+i-1}$ is either the $\alpha-$ percentile corresponding to the actual distribution of the estimated standardized residuals, $\widehat{\varsigma}_{m,T+i-1} = \widehat{\varepsilon}_{T+i-1}/\widehat{\sigma}_{m,T+i-1}$, for Gaussian models (Engle 2001), or the α-percentile of the standardized Student's t with $\hat{\nu}_m$ degrees of freedom. Notice that the estimate of ν is different for each GARCH-type specification. Notice also, that this way of calculating the percentile for the case of the Gaussian models allows to account for the observed excess kurtosis not explained yet by the GARCH-type process, thus improving their performance to forecast the VaR. On the other hand, it is worth mentioning that VaR forecasts from Student's t models were also calculated by using the percentiles of the actual distribution of $\widehat{\varsigma}_{m,T+i-1}$. In that case the validation results showed that the models did not yield more accurate forecasts. These results are not presented for the sake of simplicity but are available from the author upon request.

6.3 VaR models evaluation

In this section, models are compared in relation to their performance for predicting VaR. The comparison was conducted using the following validity criteria.

6.3.1 Regulatory loss functions

This evaluation method consists of assigning a numerical score to the VaR forecasts. This score is calculated following the current regulatory framework; i.e., regulators observe the VaR forecasts from model (financial institution) m together with the contemporaneous portfolio returns, denoted as $\left\{\widehat{VaR}^{\alpha}_{m,T+i}, r_{T+i}\right\}_{i=1}^{N}$, and then construct a numerical score under a loss function that reflects their concern. The general form of these loss functions is,

$$C^{*}_{m,T+i}(\alpha) = \begin{cases} f\left(r_{T+i}, \widehat{VaR}^{\alpha}_{m,T+i}\right) & \text{if } \; r_{T+i} < \widehat{VaR}^{\alpha}_{m,T+i} \\ g\left(r_{T+i}, \widehat{VaR}^{\alpha}_{m,T+i}\right) & \text{if } \; r_{T+i} \geq \widehat{VaR}^{\alpha}_{m,T+i}, \end{cases}$$

where $f\left(r_{T+i}, \widehat{VaR}^{\alpha}_{m,T+i}\right) \geq g\left(r_{T+i}, \widehat{VaR}^{\alpha}_{m,T+i}\right)$. Numerical scores are generated for individual VaR forecasts from every model m and confidence level α , $C^{*}_{m}(\alpha)$, for the complete regulatory sample,[11]

$$C^{*}_{m}(\alpha) = \sum_{i=1}^{N} C^{*}_{m,T+i}(\alpha). \tag{30}$$

[11] It is worth noting that in this paper we consider an out-of-sample period of 900 observations while regulators consider only the last 250 reported VaR forecasts.

Under very general conditions (cf. Diebold et al. 1998) more accurate VaR forecasts generate lower numerical scores.

Loss function that addresses the number of exceptions This function, denoted as $C_m^N(\alpha)$, only considers the number of exceptions,

$$C_{m,T+i}^N(\alpha) = \begin{cases} 1 & \text{if} \quad r_{T+i} < \widehat{VaR}_{m,T+i}^{\alpha} \\ 0 & \text{if} \quad r_{T+i} \geq \widehat{VaR}_{m,T+i}^{\alpha}. \end{cases} \tag{31}$$

According to this criterion, a correctly specified model for VaR should provide an accurate unconditional coverage probability, $\widehat{\alpha} = C_m^N(\alpha)/N$. For instance, for regulatory purposes ($\alpha = 0.01$) and for the full sample of 900 VaR forecasts, the preferred model would yield the closest to 9 exceptions, i.e., the closest $\widehat{\alpha}$ to 0.01.

Loss function that addresses the magnitude of the exceptions This type of loss functions, which we denote as $C_m^M(\alpha)$, embodies the magnitude of the exception, so providing useful information to discriminate among similar performing models in terms of the unconditional coverage probability criteria. Although several formulations are possible, we consider the following,

$$C_{m,T+i}^M(\alpha) = \begin{cases} \left(|r_{T+i}| - \left|\widehat{VaR}_{m,T+i}^{\alpha}\right|\right)^2 & \text{if } r_{T+i} < \widehat{VaR}_{m,T+i}^{\alpha} \\ 0 & \text{if } r_{T+i} \geq \widehat{VaR}_{m,T+i}^{\alpha}. \end{cases} \tag{32}$$

6.3.2 Dynamic quantile test

The dynamic quantile test, also called HIT test, of Engle and Manganelli (2004) consists of testing the joint significativity of the parameters in the following artificial regression,

$$HIT^{\alpha} = X^{\alpha}\pi + \varpi \tag{33}$$

where $\pi \in \mathbb{R}^{r+2}$, the series of HITs is defined as, $\mathbf{1}_{[r_{T+i} < VaR_{T+i}^{\alpha}]} - \widehat{\alpha}$, and the terms in the $N \times (r+2)$ matrix X are measurable$-\Omega_{t-1}$, including a column of ones, lagged HITs up to lag r, and the predicted VaR. Under the null of a correctly specified model for VaR, the HITs should be independent on everything in Ω_{t-1}. Hence, what we want to test is the null hypothesis $H_0 : \pi = 0$. The asymptotic distribution of $\widehat{\pi}_{OLS}$ under the null is given by,

$$\widehat{\pi}_{OLS} \overset{a}{\sim} N\left(0, \alpha(1-\alpha)(X'X)^{-1}\right). \tag{34}$$

Therefore, the dynamic quantile test statistic can be derived as,

$$\frac{\widehat{\pi}'_{OLS} X' X \widehat{\pi}_{OLS}}{\alpha(1-\alpha)} \overset{a}{\sim} \chi^2_{r+1}. \tag{35}$$

6.3.3 Predictive quantile loss function

The predictive quantile loss (PQL) function, $Q(\alpha)$, based on Koenker and Bassett (1978) "check" function, is a rather intuitive loss criteria that assigns $(1-\alpha)$ weight when $r_t < \widehat{VaR}_t^{\alpha}$ and α on the contrary. Thus the model that gives the lowest $Q(\alpha)$ is the preferred model for each α level.

$$Q_m(\alpha) = \frac{1}{N} \sum_{t=T+1}^{T+N} \left[\alpha - \mathbf{1}(r_t < \widehat{VaR}_{m,t}^{\alpha})\right] (r_t - \widehat{VaR}_{m,t}^{\alpha}). \tag{36}$$

This function may be regarded as "predictive" quasi-likelihood, as discussed in Komunjer (2005).

6.3.4 Results

Table 6 below reports the statistics for the evaluation of the VaR forecasts accuracy. A first observation that emerges from this table is that, according to the probability of unconditional coverage or number of exceptions, the Student's t models generally provide better forecasting performance the lower the target quantile. Although it is difficult to discriminate among Student's t models under this criterion, it is worth mentioning that the FIAPARCH-t provides the lowest number of exceptions for regulatory VaR.

Regarding the magnitude of the exception statistic, we observe that Student's t models clearly provide more accurate forecasts in relation to their Gaussian counterparts the lower the target quantile. In particular, among Gaussian models, VaR forecasts from FIAPARCH are identified as the most accurate, and the ones from GARCH and EWMA models as the least accurate. Similar conclusions can be drawn for the Student's t models, so the FIAPARCH-t and the EWMA-n are identified as the most and the least recommendable models for 2.5 and 1% VaR, under this criteria. Figure 3 below provides a illustration of the VaR forecasts from these latter two models.

According to the HIT test, it is noted that the null is more likely to be accepted the lower the quantile roughly for all models and a higher percentage of acceptance is observed for FIGARCH and FIAPARCH models. The VaR models based on the Student's t distribution perform slightly better the lower the target quantile.

Regarding the predictive quantile loss function, the results show very small differences among models and across distributions, and the criteria does not show enough power to discriminate among Student's t models are their Gaussian counterparts. Although, it is still possible to identify the VaR forecasts from the EWMA and the FIAPARCH models as the least and the most accurate in relation to this criterion.

In summary, the results in Table 6 provide evidence of that long-memory combined with asymmetries, powers and heavy-tailed distributions does help to enhance the performance of GARCH-type models to forecast VaR. Furthermore, the more stylized

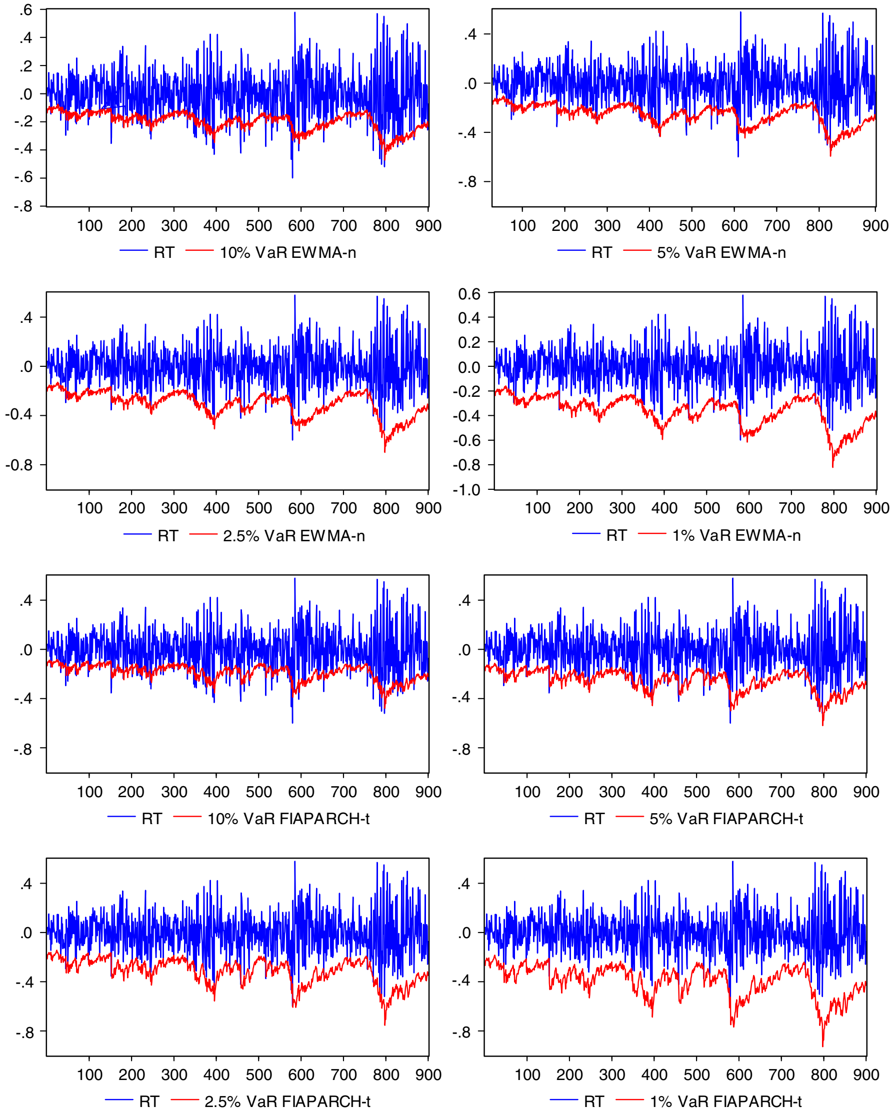

Fig. 3 Plots of (10, 5, 2.5, 1%) VaR forecasts from FIAPARCH-t and EWMA-n models against out-of-sample returns. 900 forecasts

features the GARCH model embodies the more likely are its VaR forecasts to be more accurate.[12]

[12] These findings could be generalized not only to other stock indexes rather than the IBEX-35 but also to other asset returns as much as they present the stylized features explained by the models considered in the paper (long memory, leverage effects, etc.). In particular, they are more likely to hold for exchange rates data since they present leverage effects and a high degree of persistence in volatility, as has been shown in several studies. Much less is known on long-memory in the volatility of interest rates data, so further research in this line seems worthwhile.

Table 6 VaR forecast accuracy evaluation

	Gaussian					Student's t			
	VaR^{α}	$\widehat{\alpha}$	$C^M(\alpha)$	p-v	$Q(\alpha)$	$\widehat{\alpha}$	$C^M(\alpha)$	p-v	$Q(\alpha)$
GARCH	$VaR^{.1}$	0.130	1.154	0.018	0.0282	0.138	1.203	0.003	0.0283
	$VaR^{.05}$	0.070	0.482	0.009	0.0165	0.068	0.454	0.005	0.0166
	$VaR^{.025}$	0.043	0.218	0.000	0.0093	0.031	0.176	0.005	0.0094
	$VaR^{.01}$	0.019	0.089	0.000	0.0043	0.011	0.051	0.237	0.0044
AGARCH	$VaR^{.1}$	0.128	1.139	0.003	0.0281	0.134	1.183	0.003	0.0282
	$VaR^{.05}$	0.071	0.450	0.011	0.0164	0.069	0.421	0.010	0.0165
	$VaR^{.025}$	0.042	0.177	0.002	0.0092	0.032	0.143	0.063	0.0092
	$VaR^{.01}$	0.016	0.060	0.039	0.0041	0.010	0.038	0.856	0.0043
APARCH	$VaR^{.1}$	0.127	1.116	0.045	0.0280	0.148	1.509	0.000	0.0288
	$VaR^{.05}$	0.070	0.439	0.019	0.0164	0.078	0.524	0.000	0.0167
	$VaR^{.025}$	0.038	0.171	0.027	0.0091	0.033	0.150	0.262	0.0092
	$VaR^{.01}$	0.014	0.057	0.147	0.0041	0.008	0.029	0.999	0.0043
EWMA	$VaR^{.1}$	0.118	1.066	0.072	0.0281	0.138	1.218	0.002	0.0283
	$VaR^{.05}$	0.061	0.463	0.041	0.0166	0.069	0.485	0.002	0.0167
	$VaR^{.025}$	0.033	0.226	0.106	0.0095	0.031	0.201	0.027	0.0095
	$VaR^{.01}$	0.017	0.102	0.000	0.0046	0.011	0.065	0.204	0.0045
FIGARCH	$VaR^{.1}$	0.118	1.044	0.263	0.0280	0.130	1.157	0.027	0.0281
	$VaR^{.05}$	0.057	0.422	0.023	0.0164	0.066	0.418	0.008	0.0165
	$VaR^{.025}$	0.032	0.179	0.191	0.0093	0.028	0.144	0.258	0.0092
	$VaR^{.01}$	0.011	0.070	0.000	0.0043	0.009	0.036	0.813	0.0043
FIAPARCH	$VaR^{.1}$	0.119	0.981	0.224	0.0278	0.128	1.158	0.027	0.0281
	$VaR^{.05}$	0.061	0.364	0.081	0.0162	0.067	0.397	0.026	0.0164
	$VaR^{.025}$	0.030	0.131	0.501	0.0091	0.032	0.115	0.189	0.0090
	$VaR^{.01}$	0.010	0.039	0.918	0.0040	0.007	0.026	0.999	0.0041

The table reports the empirical probability of unconditional coverage ($\widehat{\alpha}$), the magnitude of the exceptions statistic ($C^M(\alpha)$), P values for the Hit test of Engle and Manganelli (2004), denoted as p-v, and the predictive quantile loss statistic ($Q(\alpha)$), for 1 day ahead VaR forecasts from the different considered models. 900 forecasts

7 Conclusions

This paper provides a comparative empirical analysis for daily IBEX-35 returns to analyze the forecasting performance for out-of-sample volatility and VaR, of a wide range of conditional heteroscedasticity models embodying asymmetric effects, volatility powers, long-memory and heavy-tailed distributions.

The inference from the in-sample analysis gives evidence of that the goodness-of-fit of the FIAPARCH model with t distributed innovations is better than that of the rest of considered models. This is due to the fact that the FIAPARCH-t accounts for all the stylized features of stock returns volatility, unlike the other analyzed processes, which leave some of them unexplained. The assumption of Student's t -distributed innovations plays an important role since all models with t errors perform better than their Gaussian counterparts.

On the other hand, two sharp results that emerge from the out-of-sample analysis for volatility are: (a) the assumption of Student's t innovations does not contribute to get more accurate volatility forecast and, (b) asymmetric models tend to yield a better performance than models that do not allow for asymmetric responses of volatility to the sign of returns. These results are generally in line with those of Brooks and Persand (2003). Furthermore, we observe that EWMA, GARCH and FIGARCH models produce slightly but significantly the most erratic forecasts. On the other hand, the FIAPARCH models provide more accurate forecasts than the rest of the methods according to the MSPE. Besides FIAPARCH models show a higher trend to overpredict volatility in relation to other good performer models, such as, AGARCH-n and APARCH-n models.

As regards the models comparison according to their ability for VaR forecasting, the validation criteria of regulatory loss functions and HITs test show that overall the Student's t models yield a better performance than their Gaussian counterparts for 1% and 2.5% VaR, and in most of the cases also for 5% and 10% VaR. The FIAPARCH-t and the EWMA-n models being the most and the least accurate, respectively, for regulatory VaR (1% VaR). On the other hand, according to the PQL functions we observe that is very difficult to discriminate across models, although the FIAPARCH processes present the lowest PQL values.

These empirical findings generally contrast with those in prior papers suggesting that simple misspecified models provide appropriate VaR forecasts in relation to more complex ones (see for instance Lucas 2000). However, long memory specifications have not been much analyzed in a VaR forecasting context yet and, without considering those models, our results are generally consistent with those found in previous papers, such as, López and Walter (2001) and Ferreira and López (2005).

It is worth noting that a consensus has not been achieved yet in the literature regarding the best model for forecasting financial risk. This is due to the fact that the model selection process depends on so many factors, including: the data set used, the models and evaluation criteria considered, the estimator used for the unobservable volatility variable and, the final purpose of the model. See Poon and Granger (2003) for a comprehensive coverage of the status of this topic.

A conclusion that can be extracted from our results and the discussion above is not that one should use the FIAPARCH-t model and ignore the others, as it is likely that a mixture of forecasts from suitable models with optimal weights may be superior, as suggested by Poon and Granger (2003), among others. In our opinion, the FIAPARCH model and the Student's t distribution (or another suitable heavy-tailed distribution) should be considered when deciding on the models to include in the pool. Further research in this line seems worthwhile.

Acknowledgments I would like to thank Eva Ferreira, Gabriele Fiorentini, Giampiero Gallo, Ángel León, Juan Mora, Andrés Romeu, Antonio Rubia, Esther Ruiz, three anonymous referees and audiences at Alicante University and XXVII Symposium of Economic Analysis (Salamanca, 2002) for very useful comments. Financial support from the IVIE and the Spanish Ministry of Science and Technology through projects SEJ2006/06104 and BEC2002-03797 is gratefully acknowledged. The paper was finished while the author visited the London School of Economics and Political Science, whose hospitality is gratefully acknowledged. Responsibility for the contents remains solely with the author.

Appendix A. Nested processes for the conditional variance

The FIAPARCH(p, d, q) of Tse (1998) includes the following specifications as special cases:

1. The FIPARCH(p, d, q) process when $\gamma = 0$,

$$\sigma_t^\delta = \frac{\omega}{1-\beta(1)} + \left[1 - [(1-\beta(L)]^{-1}\phi(L)(1-L)^d\right] |\varepsilon_t|^\delta . \quad \text{(A.1)}$$

2. The FIAGARCH(p, d, q) model when $\delta = 2$,

$$\sigma_t^2 = \frac{\omega}{1-\beta(1)} + \left[1 - [(1-\beta(L)]^{-1}\phi(L)(1-L)^d\right] (|\varepsilon_t| - \gamma\varepsilon_t)^2. \quad \text{(A.2)}$$

3. The FIGARCH(p, d, q) model of Baillie et al. (1996), Eq. (11), when $\delta = 2$ and $\gamma = 0$.
4. The IGARCH(p, q) model of Engle and Bollerslev (1986), Eq. (9), when $d = 1$, $\delta = 2$ and $\gamma = 0$.
5. The APARCH(p, q) model of Ding et al. (1993), Eq. (14), when $d = 0$, and the other eight specifications included in it, namely:
   - ARCH(p) model of Engle (1982), Eq. (6), when $\delta = 2$, $\gamma_k = 0 \forall k$, and $\beta_s = 0\,\forall s$.
   - GARCH(p, q) model of Bollerslev (1986), Eq. (7), when $\delta = 2$ and $\gamma_k = 0\,\forall k$.
   - PARCH(p, q) model, when $d = 0$ and $\gamma_k = 0\,\forall k$,

$$\sigma_t^\delta = \omega + \sum_{k=1}^{p} \alpha_k\, |\varepsilon_{t-k}|^\delta + \sum_{s=1}^{q} \beta_s \sigma_{t-s}^\delta . \quad \text{(A.3)}$$

   - AGARCH(p, q) model of Glosten et al. (1993), when $\delta = 2$,

$$\sigma_t^2 = \omega + \sum_{k=1}^{p} \alpha_k (|\varepsilon_{t-k}| - \gamma_k \varepsilon_{t-k})^2 + \sum_{s=1}^{q} \beta_s \sigma_{t-s}^2 , \quad \text{(A.4)}$$

   - Absolute Value GARCH of orders p and q, AVGARCH(p, q), model of Taylor/Schwert (1989) when $\delta = 1$ and $\gamma_k = 0\,\forall k$,

$$\sigma_t = \omega + \sum_{k=1}^{p} \alpha_k\, |\varepsilon_{t-k}| + \sum_{s=1}^{q} \beta_s \sigma_{t-s} , \quad \text{(A.5)}$$

– Threshold ARCH of order p, TARCH(p), model of Zakoian (1994) when $\delta = 1$ and $\beta_s = 0\ \forall s$,

$$\sigma_t = \omega + \sum_{k=1}^{p} \alpha_k^{+} \varepsilon_{t-k}^{+} - \sum_{k=1}^{p} \alpha_k^{-} \varepsilon_{t-k}^{-}, \tag{A.6}$$

where $\alpha_k^{+} = \alpha_k(1-\gamma_k)$, $\alpha_k^{-} = \alpha_k(1+\gamma_k)$ and $\varepsilon_{t-k}^{-} = \varepsilon_{t-k} - \varepsilon_{t-k}^{+}$ with,

$$\varepsilon_{t-k}^{+} = \begin{Bmatrix} \varepsilon_{t-k} & \text{if } \varepsilon_{t-k} > 0 \\ 0 & \text{otherwise} \end{Bmatrix}. \tag{A.7}$$

– Non-linear ARCH of order p, NARCH(p), model of Higgins and Bera (1992) when $\gamma_k = 0\ \forall k$ and $\beta_s = 0\ \forall s$,

$$\sigma_t^{\delta} = \omega + \sum_{k=1}^{p} \alpha_k \left|\varepsilon_{t-k}\right|^{\delta}, \tag{A.8}$$

– log-GARCH(p) model of Pantula (1986), which is the limiting case of the APARCH(p, q) when $\delta \to 0$, $\gamma_k = 0\ \forall k$ and $\beta_s = 0\ \forall s$,

$$\log \sigma_t = \left(1 - \sum_{k=1}^{p} \alpha_k\right) \log \varpi - \sum_{k=1}^{p} \alpha_k \log \sqrt{2/\pi} + \sum_{k=1}^{p} \alpha_k \log \left|\varepsilon_{t-k}\right|, \tag{A.9}$$

where

$$\log \varpi = \lim_{\delta \to 0} \frac{\varpi^{\delta} - 1}{\delta}, \tag{A.10}$$

and ϖ^{δ} is the unconditional expectation of σ_t^{δ},

$$\mathbf{E}(\sigma_t^{\delta}) = \varpi^{\delta} = \frac{\omega}{1 - \sum_{k=1}^{p} \alpha_k \mathbf{E}(|\eta_{t-k}| - \gamma_k \eta_{t-k})^{\delta}}. \tag{A.11}$$

The most general form of quadratic ARCH-type processes, i.e., the generalized quadratic ARCH (G-QARCH) model of Sentana (1995) is not nested in the FIAPARCH(p, d, q) structure. The G-QARCH(p, q) model is defined by

$$\sigma_t^2 = \omega + \sum_{k=1}^{p} \gamma_k \varepsilon_{t-k} + \sum_{k=1}^{p} \alpha_k \varepsilon_{t-k}^2 + \sum_{s=1}^{q} \beta_s \sigma_{t-s}^2. \tag{A.12}$$

and encompasses the GARCH(p, q) when the asymmetry parameters $\gamma_k = 0\ \forall k$.

The GARCH-in-Mean, GARCH-M, process introduced by Engle et al. (1987) to account for the possible effect of the expected asset risk on the expected return, is nested in the FIAPARCH process as long as we specify the corresponding conditional mean equation. A variable r_t is said to follow a GARCH-M(p, q) model if

$$r_t = \mu + \boldsymbol{\zeta}'\mathbf{X}_t + \delta\sigma_t^2 + \varepsilon_t, \quad \varepsilon_t = \sigma_t\eta_t,$$
$$\sigma_t^2 = \omega + \sum_{k=1}^{p} \alpha_k \varepsilon_{t-k}^2 + \sum_{s=1}^{q} \beta_s \sigma_{t-s}^2. \tag{36}$$

where $\boldsymbol{\zeta}$ and $\mathbf{X}_t$ are $\kappa \times 1$ vectors of parameters and, exogenous and lagged dependent variables, respectively, and the coefficient δ is a measure of the risk-return trade-off. A preliminary specification test for our IBEX-35 returns sample, revealed the non-significativity of the GARCH-M term and that a MA(1) process is appropriate for the conditional mean.[13] It is worth noting that although it is known that a GARCH-M specification would present unique estimation and testing problems in our two-stage estimation procedure, because of the explicit dependence of the conditional mean on the conditional variance, this need not be an obstacle for VaR forecasting purposes, for which GARCH-M models are likely to provide a good out-of-sample performance if risk-return trade-off effects are present in-sample.

It is also worth mentioning that alternative GARCH specifications to model asymmetries defined in terms of the logarithm of the conditional variance, such as, the FIEGARCH model of Bollerslev and Mikkelsen (1996), and the exponential GARCH (EGARCH) model of Nelson (1991) are not nested in the FIAPARCH structure.

More recently, Davidson (2004) has proposed an hyperbolic GARCH (HYGARCH) process that generalizes the FIGARCH, and allows for a (possible) faster non-geometric rate of decay of its CIRW for which weakly stationarity and long-run dependence are possible, likewise the FIEGARCH model, although the HYGARCH allows for non-stationarity as well; see Ñíguez and Rubia (2006) for an application of this process to forecast the covariance matrix of an exchange-rates portfolio. Extensions of the HYGARCH process to include asymmetry and powers constitute interesting avenues for future research.

Appendix B. Notes on the test of Diebold and Mariano (1995)

For a given loss function of the forecast errors, $L(e_{T+i})$ $(i = 1, \ldots, N)$, e.g., MAPE $|e_{T+i}| = \left|\widehat{\varepsilon}_{T+i}^2 - \widehat{\sigma}_{T+i}^2\right|$, the DM null hypothesis of equal predictive ability of forecasts from two models x and y is,

$$H_0 : E\left[d_{T+i}\right] = 0, \tag{B1}$$

where

$$d_{T+i} = L(e_{x,T+i}) - L(e_{y,T+i}). \tag{B2}$$

[13] These results are not presented for the sake of symplicity but they are available from the author upon request.

The DM test statistic is of the form,

$$DM = \frac{\overline{d}}{\sqrt{2\pi \widehat{f_d}(0)/N}}, \tag{B3}$$

where $\overline{d} = \frac{1}{N}\sum_{i=1}^{N} d_{T+i}$, and $\widehat{f_d}(0)$ is a consistent estimate of the spectral density function of the loss differential at frequency 0. This statistic is asymptotically distributed as a standard normal under the null.

If the forecast is optimal, the K step ahead forecast errors are at most independent of all previous ones. So, consistent estimates for $f_d(0)$ can be obtained by using the HAC Newey–West (1987) estimator with bandwidth parameter $K - 1$.

References

Andersen TG, Bollerslev T (1998) Answering the skeptics: yes, standard volatility models do provide accurate forecasts. Int Econ Rev 39:885–905

Arteche J (2004) Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models. J Econom 119:131–154

Baillie RT (1996) Long memory processes and fractional integration in econometrics. J Econom 73:5–59

Baillie RT, Bollerslev T, Mikkelsen H (1996) Fractionally integrated generalized autoregressive conditional heteroscedasticity. J Econom 74:3–30

Bauwens L, Laurent S, Rombouts JVK (2006) Multivariate GARCH models: a survey. J Appl Econom 21:79–109

Black F (1976) Studies of stock price volatility changes. In: Proceedings from the American statistical association. Business and economic statistics section, pp 177–181

Bollerslev T (1986) Generalized autoregressive conditional heteroscedasticity. J Econom 31:307–327

Bollerslev T (1987) A conditional heteroskedastic time series model for speculative prices and rates of return. Rev Econ Stat 69:542–547

Bollerslev T, Chou RY, Kroner KF (1992) ARCH modeling in finance. J Econom 52:5–59

Bollerslev T, Engle RF (1993) Common persistence in conditional variances. Econometrica 61:167–186

Bollerslev T, Mikkelsen H (1996) Modeling and pricing long memory in stock market volatility. J Econom 73:151–184

Bollerslev T, Mikkelsen H (1999) Long-term equity anticipation securities and stock market volatility dynamics. J Econom 92:75–99

Bollerslev T, Wooldridge JM (1992) Quasi-maximum likelihood estimation and inference in dynamics models with time-varying covariances. Econom Rev 11:143–172

Brailsford TJ, Faff RW (1996) An evaluation of volatility forecasting techniques. J Bank Finance 20:419–438

Breidt FJ, Crato N, De Lima P (1998) The detection and estimation of long memory in stochastic volatility. J Econom 83:325–348

Brooks C, Persand G (2003) The effect of asymmetries on stock index return value-at-risk estimates. J Risk Finance 4:29–42

Caporin M (2003) Identification of long memory GARCH models. Stat Methods Appl 12:133–151

Chung CF (1999) Estimating the fractionally integrated GARCH model. Manuscript. National Taiwan University, Taiwan

Conrad C, Haag BR (2006) Inequality constraints in the fractionally integrated GARCH model. J Financ Econom 4:413–449

Danielsson J, Shin H, Zigrand JP (2004) The impact of risk regulation on price dynamics. J Bank Finance 28:1069–1087

Davidson J (2004) Moment and memory properties of linear conditional heteroscedasticity models, and a new model. J Bus Econ Stat 22:16–29

Diebold FX, Inoue A (1999) Long memory and structural change. Manuscript. New York University, Stern School of Business, Department of Finance
Diebold FX, Mariano R (1995) Comparing predictive accuracy. J Bus Econ Stat 13:253–264
Diebold FX, Gunther TA, Tay SA (1998) Evaluating density forecasts with applications to financial risk management. Int Econ Rev 39:863–883
Ding Z, Granger CWJ (1996) Modeling volatility persistence of speculative returns: a new approach. J Econom 73:185–215
Ding Z, Granger CWJ, Engle RF (1993) A long memory property of stock market returns and a new model. J Empir Finance 1:83–106
Engle RF (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation. Econometrica 50:987–1008
Engle RF (2001) Garch 101: the use of ARCH/GARCH models in applied econometrics. J Econ Perspect 15:157–168
Engle RF, Bollerslev T (1986) Modeling the persistence of conditional variances. Econom Rev 5:1–50
Engle RF, Lilien DM, Robins RP (1987) Estimating time varying risk premia in the term structure: the ARCH-M model. Econometrica 55:391–407
Engle RF, Manganelli S (2004) CAViaR: conditional value at risk by regression quantiles. J Bus Econ Stat 22:367–381
Ferreira MA, López JA (2005) Evaluating interest rate covariance models within a Value-at-Risk framework. J Financ Econom 3:126–168
Fiorentini G, Sentana E, Calzolari G (2003) Maximum likelihood estimation and inference in multivariate conditional heteroskedastic dynamic regression models with Student t innovations. J Bus Econ Stat 21:532–546
Giraitis L, Kokoszka P, Leipus R (2000) Stationary ARCH models: dependence structure and central limit theorem. Econom Theory 16:3–22
Glosten LR, Jagannathan R, Runkle D (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. J Finance 48:1779–1801
Gradshteyn IS, Ryzhnik IM (1980) Tables of integrals, series and products. Academic, New York
Granger CWJ, Ding Z (1996) Varieties of long memory models. J Econom 73:61–77
Hansen PR, Lunde A (2005) A forecast comparison of volatility models: does anything beat a GARCH(1,1)? J Appl Econom 20:873–889
Hansen PR, Lunde A (2006) Consistent ranking of volatility models. J Econom 131:97–121
Hidalgo J (1997) Non-parametric estimation with strongly dependent multivariate time-series. J Time Ser Anal 18:95–122
Higgins M, Bera A (1992) A class of non-linear ARCH models. Int Econ Rev 33:137–158
Hosking JRM (1981) Fractional differencing. Biometrika 68:165–176
Morgan JP (1996) RiskMetrics—technical document, 4th edn. Morgan Guaranty Trust Company, New York
Jorion P (2001) Value at risk: the new benchmark for controlling financial risk. McGraw-Hill, London
Jorion P (2002) Fallacies about the effects of market risk management systems. J Risk 5:75–96
Koenker R, Bassett G (1978) Regression quantiles. Econometrica 46:33–50
Komunjer I (2005) Quasi-maximum likelihood estimation for conditional quantiles. J Econom 128:137–164
Lee KY (1991) Are the GARCH models best in out-of-sample performance? Econ Lett 37:305–308
León A, Mora J (1999) Modeling conditional heteroscedasticity: application to the IBEX 35 stock-return index. Span Econ Rev 3:215–238
Ling S, McAleer M (2002) Necessary and sufficient moment conditions for the GARCH(r, s) and asymmetric power GARCH(r, s) models. Econom Theory 18:722–729
Lobato IN, Robinson PM (1997) A non-parametric test for I(0). Rev Econ Stud 65:475–495
Lobato IN, Savin NE (1998) Real and spurious long memory properties of stock market data. J Bus Econ Stat 16:261–268
López JA (1999) Methods for evaluating value-at-Risk estimates. Federal Reserve Bank of San Francisco Economic Review, vol 2, pp 3–17
López JA, Walter CA (2001) Evaluating covariance matrix forecasts in a value-at-risk framework. J Risk 3:69–98
Lucas A (2000) A note on optimal estimation from a risk-management perspective under possibly misspecified tail behavior. J Bus Econ Stat 18:31–39
Mandelbrot B (1963) The variation of certain speculative prices. J Bus 36:1279–1313

Mármol F, Reboredo JC (2000) Dependencia temporal en la volatilidad estocástica. Una aplicación al índice IBEX 35. Centro de Investigación Económica y Financiera Working Paper no. 4. Fundación Caixa Galicia
Mincer J, Zarnowitz V (1969) The evaluation of economic forecasts. In: Mincer J (ed) Economic forecasts and expectations. NBER, New York
Morana C, Beltratti A (2004) Structural change and long-range dependence in volatility of exchange rates: either, neither or both? J Empir Finance 11:629–658
Muller UA, Dacorogna MM, Davé RD, Olsen RB, Pictet OV, von Weizsacker JE (1997) Volatilites of different times resolutions—analyzing the dynamics of markets components. J Empir Finance 4:213–239
Nelson CR (1976) Applied time series analysis. Holden-Day, San Francisco
Nelson DB (1991) Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59:347–370
Newey W, West K (1987) A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55:703–708
Newey W, Steigerwald D (1997) Asymptotic bias for quai-maximum-likelihood estimators in conditional heteroskedasticity models. Econometrica 65:587–599
Ñíguez TM, Rubia A (2006) Forecasting the conditional covariance matrix of a portfolio under long-run temporal dependence. J Forecast 25:439–458
Pagan A (1986) Two stage and related estimators and their applications. Rev Econ Stud 53:517–538
Pantula SG (1986) Modeling the persistence of conditional variance: a comment. Econom Rev 5:71–74
Perron PY, Ng S (1996) Useful modifications to some unit-root tests with dependent errors and their local asymptotic properties. Rev Econ Stud 63:435–463
Poon SH, Granger CWJ (2003) Forecasting volatility in financial markets. J Econ Lit 41:478–539
Robinson PM (2003) Time series with long memory. Oxford University Press, Oxford
Ryden T, Teräsvirta T, Asbrink S (1998) Stylized facts of daily returns series and the hidden Markov model. J Appl Econom 13:217–244
Sentana E (1995) Quadratic ARCH models. Rev Econ Stud 62:639–661
Sentana E, Wadhwani S (1992) Feedback traders and stock return autocorrelations: evidence from a century of daily data. Econ J 102:415–425
Schwert GW (1989) Why does stock market volatility change over time? J Finance 44:1115–1153
Straumann D, Mikosch T (2006) Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: a stochastic recurrence equations approach. Ann Stat 34:2449–2495
Tse YK (1998) The conditional heteroscedasticity of the yen-dollar exchange rate. J Appl Econom 13:49–55
Vilasuso J (2002) Forecasting exchange rate volatility. Econ Lett 76:59–64
Weiss AA (1986) Asymptotic theory of ARCH models: estimation and testing. Econom Theory 2:107–131
White H (1982) Maximum likelihood estimation of misspecified models. Econometrica 50:1–26
White H (1994) Estimation, inference and specification analysis. Cambridge University Press, Cambridge
Zakoian J (1994) Threshold heteroscedasticity model. J Econ Dyn Control 15:931–955