Persistence and Kurtosis in GARCH and Stochastic Volatility Models

M. Angeles Carnero
Universidad de Alicante

Daniel Peña
Universidad Carlos III de Madrid

Esther Ruiz
Universidad Carlos III de Madrid

ABSTRACT

This article shows that the relationship between kurtosis, persistence of shocks to volatility, and first-order autocorrelation of squares is different in GARCH and ARSV models. This difference can explain why, when these models are fitted to the same series, the persistence estimated is usually higher in GARCH than in ARSV models, and, why gaussian ARSV models seem to be adequate, whereas GARCH models often require leptokurtic conditional distributions. We also show that introducing the asymmetric response of volatility to positive and negative returns does not change the conclusions. These results are illustrated with the analysis of daily financial returns.

KEYWORDS: ARSV, EGARCH, leverage effect, QGARCH

Financial returns observed, for example, on a daily or weekly basis, are mainly characterized by high kurtosis, small first-order autocorrelation of squares, and slow decay of the autocorrelations of squares toward zero. These three characteristics have been documented by a large number of authors [see, e.g., Liesenfeld and Jung (2000) and Loudon, Watt, and Yadav (2000), among many others]. As an illustration, Table 1 reports the sample moments for four series of daily financial

Table 1 Descriptive statistics of daily returns.

Series	US-CAN	US-JAP	BOMBAY	S & P 500
T	1963	1963	1262	2888
Mean	0.0084	-0.0077	0.0111	0.0575
S.D.	0.3094	0.7796	1.8471	0.9085
Skewness	-0.1449^*	-0.5858^*	-0.1329	-0.6108^*
Kurtosis	5.6080*	7.3970*	6.8799*	9.6405*
Autocorrelations of squared returns				
$r_2(1)$	0.1200*	0.2500*	0.1300*	0.1700*
$r_2(2)$	0.0900*	0.1300*	0.1300*	0.0900*
$r_2(5)$	0.1400*	0.1100*	0.1000*	0.1500*
$r_2(10)$	0.0800*	0.0500*	0.0400*	0.0700*
$Q_2(20)$	380*	377*	146*	497*

T: Sample size.
$r_2(\tau)$: Order τ autocorrelation of squares y_t^2.
$Q_2(20)$: Box-Ljung statistic for y_t^2.
*Significant at the 5% level.

returns. These series are returns of the exchange rates of the US dollar against the Canadian dollar (US-CAN) and the Japanese yen (US-JAP) from January 1993 until October 2000, and of the Bombay Stock Market (BOMBAY) and the Standard and Poor's (S&P 500) indexes from January 1995 to October 2000 and from November 1987 to December 1998, respectively.[1] Before their analysis the series were filtered to get rid of some weak dynamics in the conditional mean. Table 1 shows that the values of the first-order autocorrelation of squared returns are rather small, ranging from 0.12 to 0.25. On the other hand, the kurtosis is always greater than 3, with the smallest kurtosis corresponding to the US-CAN returns (5.61) and the largest to the S&P 500 (9.64). These series are plotted in Figure 1, together with a kernel estimate of their densities and their sample autocorrelation functions of the squares. The series plots show volatility clustering, which often characterizes high-frequency time series of returns. Also notice that the autocorrelations of squares decay slowly toward zero, being significant for very long lags, although their magnitudes are rather small.

A simple model able to generate these effects specifies the returns, y_t, as the product of two processes, $y_t = \varepsilon_t \sigma_t$, where ε_t is a serially independent and identically distributed process with zero mean, unit variance, and finite fourth-order moment, independent of the second process, σ_t, which is known as volatility in the financial literature. A wide spectrum of models have been proposed for σ_t and the most often used are the generalized autoregressive conditional heteroscedasticity (GARCH) model, introduced independently by Bollerslev (1986) and

---

[1] The series were obtained from the web page http://pacific.commerce.ubc.ca/xr/ (Prof. Werner Antweiler, University of British Columbia, Vancouver, Canada).
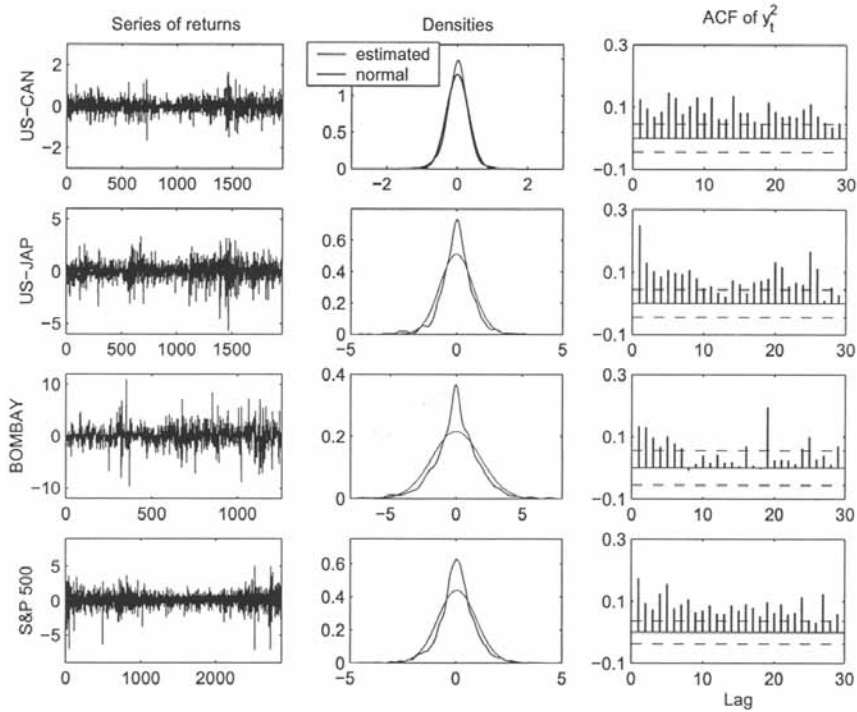
Figure 1 Financial time-series observed daily.

Taylor (1986) generalizing the model proposed by Engle (1982), and the auto-regressive stochastic volatility (ARSV) model, proposed by Taylor (1986).

Two characteristics are often observed when both GARCH and ARSV models are fitted to the same series of returns. First, the persistence of volatility implied by GARCH(1,1) models is usually higher than that implied by ARSV(1) models. For example, Taylor (1994) found that, for daily deustche mark against the dollar exchange rate returns, the persistence parameter estimated by the GARCH(1,1) model was 0.97 while that estimated by the ARSV(1) model was 0.94. Similar results have been found by Shephard (1996), Kim, Shephard, and Chib (1998), Hafner and Herwartz (2000), and Anderson (2001). Second, the gaussianity assumption for the innovations, ε_t, is adequate for ARSV(1) models, while the noise of the GARCH(1,1) specification requires a distribution with fat tails. For instance, Shephard (1996) finds that a Student-GARCH(1,1) model has a similar diagnosis to the normal-ARSV(1). Ghysels, Harvey, Renault (1996), Kim, Shephard, and Chib (1998) and Hafner and Herwartz (2000), among others, report similar findings. Measuring the persistence of volatility correctly is important because it is often estimated very close to one, which has implications for the theoretical models used to value the price of financial stocks. On the other hand, whether the distribution of ε_t is gaussian or has fat tails has implications mainly for inference on the models fitted to represent the dynamic evolution of volatility [see, e.g., Hall and Yao (2003)].

In this article we show that, although both GARCH and ARSV models are able to explain excess kurtosis and significant autocorrelations of squares with a slow rate of decay, the relationship between the persistence of shocks to the volatility, first-order autocorrelation of squares, and kurtosis implied by each of these models is different. This difference could explain the empirical results previously mentioned. The ARSV(1) model is more flexible than the GARCH(1,1) model to simultaneously represent the values of these three coefficients usually observed in real data.

Another stylized fact often observed in high-frequency financial returns is the asymmetric response of σ_t^2 to positive and negative changes in prices [see, e.g., Campbell and Hentchel (1992) and Shephard (1996), among many others]. In this article we introduce this asymmetry in the GARCH model using the specification proposed by Sentana (1995), known as the quadratic GARCH (QGARCH) model, because its statistical properties are easy to derive. On the other hand, we consider the asymmetric ARSV model proposed by Harvey and Shephard (1996). Finally, we analyze the properties of the exponential GARCH (EGARCH) model proposed by Nelson (1991). The latter model was the first model proposed to represent the asymmetric response of volatility to positive and negative returns, and it is especially interesting because it shares some of its properties with the GARCH and ARSV models. Consequently the specification of the volatility in the EGARCH model can be considered to be between the specifications assumed by the GARCH and ARSV models. We show that introducing this asymmetry does not substantially change the relationship between the three quantities of interest in any of these asymmetric conditionally heteroscedastic models with respect to the corresponding symmetric models.

The article is organized as follows. In Section 1 we describe and compare the main properties of the GARCH(1,1) and ARSV(1) models. The results of this section are illustrated by means of several Monte Carlo experiments. In Section 2 we allow for asymmetric responses of volatility to positive and negative returns in both models. Section 3 contains an empirical application where the main results are illustrated with real data. Section 4 concludes the article.

1 PROPERTIES OF GARCH AND ARSV MODELS

1.1 The GARCH(1,1) Model

GARCH models are very popular for representing the dynamic evolution of the volatility of financial returns and have been extensively analyzed in the literature [see, e.g., Bollerslev, Engle, and Nelson (1994), Engle (1994), Bera and Higgins (1995), Diebold and López (1995), and McAleer and Oxley (2003), among many others]. If y_t follows a GARCH(1,1) model, then the volatility is given by

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2, \tag{1}$$

where ω, α, and β are parameters such that $\omega > 0$ and $\alpha, \beta \geq 0$. These positivity conditions are required to guarantee the existence of the conditional variance and
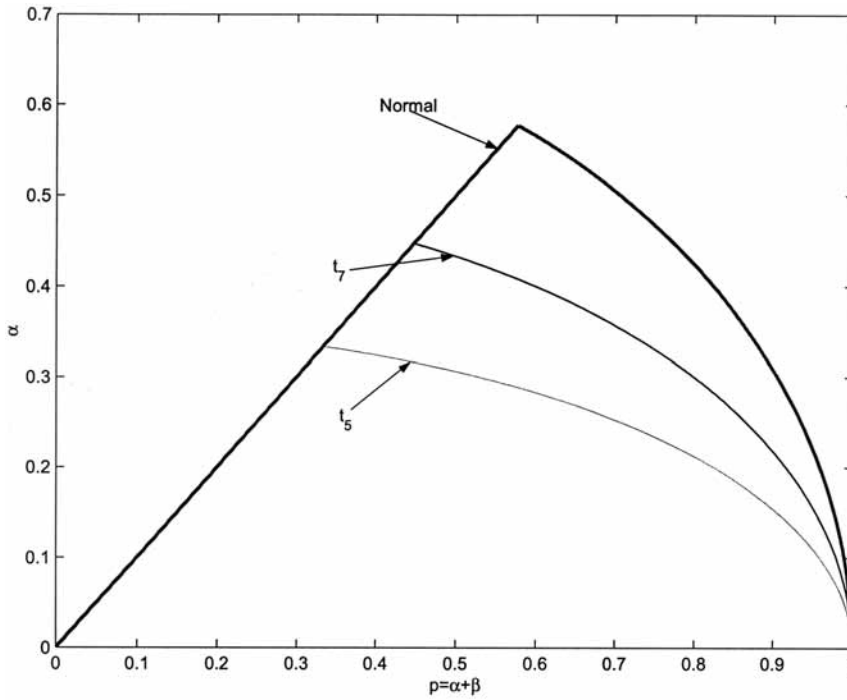
Figure 2 Areas where κ_y is finite.

to avoid degenerating the process y_t. Furthermore, if $p = \alpha + \beta < 1$, y_t has finite unconditional variance given by $\frac{\omega}{1-p}$ and it is covariance stationary. The quantity p is usually considered to be a measure of the persistence of shocks to the volatility.

A GARCH(1,1) model can be stationary without having a finite fourth-order moment. The condition for the existence of the fourth-order moment is given by

$$p^2 + \alpha^2(\kappa_\varepsilon - 1) < 1, \tag{2}$$

where κ_ε is the kurtosis of ε_t [see, e.g., He and Teräsvirta (1999)]. If this condition is satisfied, the kurtosis is given by

$$\kappa_y = \kappa_\varepsilon \left[1 - \frac{\alpha^2(\kappa_\varepsilon - 1)}{1 - p^2}\right]^{-1}. \tag{3}$$

This equation shows that, as is well known, persistence and kurtosis are highly tied up in GARCH models. Furthermore, Bai, Rusell, and Tiao (2003) show that the overall kurtosis of y_t can be decomposed into the kurtosis induced by volatility clustering and the kurtosis induced by ε_t and that the contribution of both components to the overall kurtosis is symmetric.

Figure 2 plots the area of possible values of α as a function of p defined by Equation (2) for three alternative values of κ_ε: 3, 5, and 9 corresponding to gaussian, Student-7 (t_7) and Student-5 (t_5) errors, respectively. Notice that this area is larger when ε_t is gaussian. In this case, if $p > 0.58$, the ARCH parameter, α,

should decrease as p increases. For example, if $\alpha + \beta = 0.95$, the maximum value of α is 0.22, while if $\alpha + \beta = 0.98$, the maximum value of α is 0.14. On the other hand, if the distribution of the innovations has fat tails, then the area of allowed values of α is smaller than under conditional gaussianity. For example, if the distribution of ε_t is a Student-7 distribution and $p = 0.98$, then the maximum value of α is 0.10, while if ε_t is a Student-5, the maximum value of α is 0.07. Consequently, if the persistence is high, as is often the case in real series of returns, and the conditional distribution has fat tails, the GARCH model only has finite kurtosis if the ARCH parameter is rather small.

It is important to notice that Equation (2) shows that, in GARCH(1,1) models, the dynamics of the volatility are severely restricted to guarantee that the fourth-order moment is finite. These dynamics appear in the autocorrelation function (ACF) of the squares given by

$$\rho_2(h) = \begin{cases} \dfrac{\alpha(1 - p^2 + p\alpha)}{1 - p^2 + \alpha^2}, & h = 1 \\[2ex] \rho_2(1)p^{h-1}, & h > 1 \end{cases} \tag{4}$$

[see Bollerslev (1988)]. This ACF has the same pattern as an ARMA(1,1) process with autoregressive parameter p. Therefore the autocorrelations of squares decay exponentially to zero with parameter p, which can be considered as a measure of the persistence of the volatility process. On the other hand, when $\alpha = 0$, $\rho_2(1) = 0$ and y_t is conditionally homoscedastic, there is not dependence of squared returns. If $\alpha \neq 0$, then Equation (4) shows that, for a given persistence parameter, p, the magnitude of $\rho_2(1)$ and, consequently, of successive autocorrelations, increases with α. Consider, for example, $p = 0.95$, then for $\alpha = 0.05$, 0.1, and 0.15, the corresponding $\rho_2(1)$ are 0.07, 0.18, and 0.3, respectively. Therefore, for a given persistence, α measures the dependence between squared observations and can be interpreted as the parameter leading the volatility dynamics. Finally, for a given value of α, the autocorrelations increase with the persistence. In the limit, if $p = 1$, then $\rho_2(1) = 1$. However, as we have seen before, once $p < 1$, and consequently, y_t is stationary, it is possible to have series with large persistence and small autocorrelations of squares.

From Equations (3) and (4), it is possible to derive the following relationship among κ_y, $\rho_2(1)$, and p:

$$\rho_2(1) = \frac{\sqrt{\dfrac{(\kappa_y - \kappa_\varepsilon)(1 - p^2)}{(\kappa_\varepsilon - 1)\kappa_y}}\left[1 - p^2 + p\sqrt{\dfrac{(\kappa_y - \kappa_\varepsilon)(1 - p^2)}{(\kappa_\varepsilon - 1)\kappa_y}}\right]}{1 - p^2 + \dfrac{(\kappa_y - \kappa_\varepsilon)(1 - p^2)}{(\kappa_\varepsilon - 1)\kappa_y}}. \tag{5}$$

Figure 3 represents this relationship for the same three alternative distributions of ε_t considered before. This figure also plots a box that contains the values of the kurtosis and first-order autocorrelations of squares usually encountered in empirical applications, that is, κ_y between 5 and 10 and $\rho_2(1)$ between 0.1 and 0.2. Figure 3 shows that the three surfaces are very bent toward high persistence. Therefore,
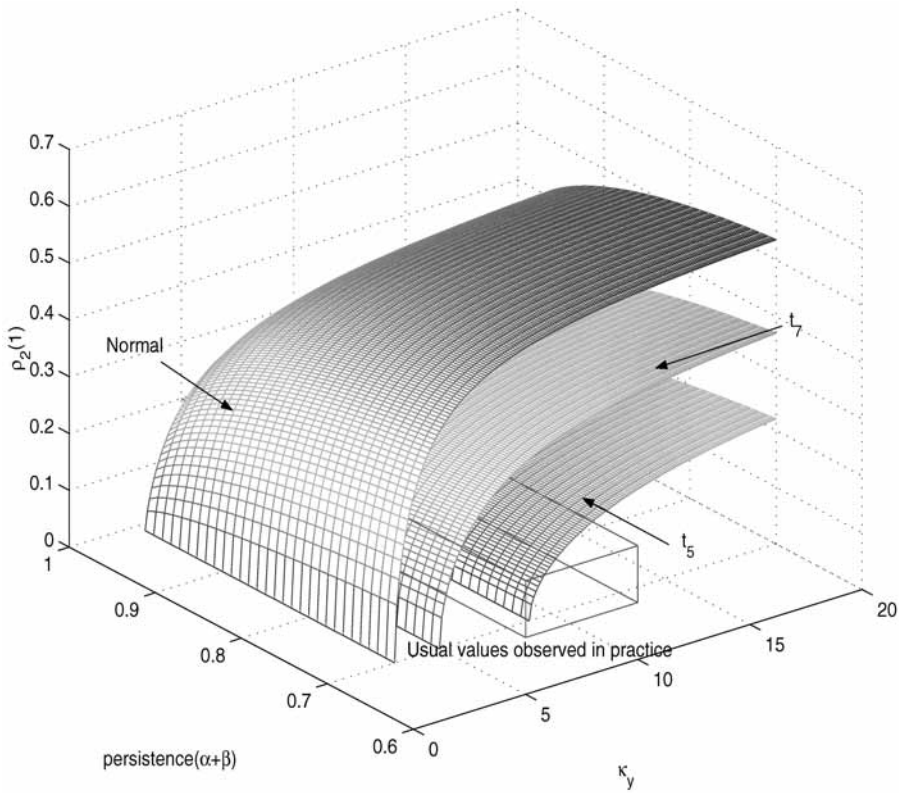
Figure 3 Relationship between kurtosis, first-order autocorrelation of squared observations, and persistence for GARCH(1,1) models.

for a given kurtosis, the first-order autocorrelation, $\rho_2(1)$, is smaller the larger the persistence. On the other hand, given the persistence, this autocorrelation increases with the kurtosis. Furthermore, it is important to point out that a normal-GARCH(1,1) model is only able to represent the values of the kurtosis and $\rho_2(1)$ within the box if the persistence is very close to one. Consequently, if the series analyzed has simultaneously high kurtosis and low first-order autocorrelation of squares, $\alpha+\beta$ could be estimated very close to one, even if the persistence of shocks to volatility is not so high. Furthermore, even if the persistence of the volatility is high, high kurtosis is associated with values of $\rho_2(1)$ larger than the values often observed in practice [see also Teräsvirta (1996) and Bai, Russell, and Tiao (2003)].

Assuming that the distribution of ε_t is a heavy-tailed distribution like, for example, the Student's t distribution, may improve the adequacy of the GARCH(1,1) model to characterize the stylized facts observed in practice. This is in agreement with Figure 3, which shows that GARCH(1,1) models with a conditional Student's t distribution are better at explaining simultaneously the three

stylized facts observed in real-time series than the gaussian alternatives. For example, consider $\alpha + \beta = 0.95$ and $\kappa_y = 10$. In this case, if ε_t were gaussian or a Student's t distribution with 7 or 5 degrees of freedom, $\rho_2(1)$ would be 0.3831, 0.2037, or 0.0462, respectively. For $\alpha + \beta = 0.99$ and $\kappa_y = 10$, if ε_t were gaussian or a Student's t distribution with 7 or 5 degrees of freedom, $\rho_2(1)$ would be 0.3185, 0.1543, or 0.0278, respectively. However, remember that in Student-GARCH models with high persistence and high but finite kurtosis, the parameter α is heavily restricted toward zero. Consequently the apparent advantage of introducing Student's t innovations in GARCH models is limited.

1.2 The ARSV(1) Model

Alternatively, the volatility can be represented by ARSV models that assume that σ_t is a latent variable. Surveys on the properties of ARSV models are given by Taylor (1994), Ghysels, Harvey, and Renault (1996), and Shephard (1996). In the simplest case, the volatility follows an autoregressive process after being transformed into logarithms. Therefore the model for volatilities, denoted ARSV(1), is given by

$$\log \sigma_t^2 = \omega + \phi \log \sigma_{t-1}^2 + \eta_t, \tag{6}$$

where η_t is assumed to be a gaussian white noise process with zero mean and variance σ_η^2, independent of ε_t. Andersen et al. (2003) show that the hypothesis of lognormality for volatility is adequate to represent the distribution of daily realized volatilities of real series of intraday returns. The restriction $|\phi| < 1$ guarantees the stationarity of y_t. The kurtosis of y_t is given by $\kappa_y = \kappa_\varepsilon \exp(\sigma_h^2)$, where $\sigma_h^2 = \sigma_\eta^2 / (1 - \phi^2)$. Notice that, in the ARSV(1) model, if κ_ε is finite, the condition for the existence of the kurtosis is the stationarity condition, that is $|\phi| < 1$. Therefore, as far as the model is stationary, the dynamic evolution of the volatility is not further restricted to guarantee the existence of the fourth-order moment. However, it is also important to notice that given the kurtosis of returns, the parameter σ_η^2, that allows the volatility to evolve over time, should decrease as the autoregressive parameter, ϕ increases. Therefore, if the kurtosis is constant, the model approaches homoscedasticity as the persistence increases [see Harvey and Streibel (1998)]. Finally, Bai, Russell, and Tiao (2003) show that, as in the GARCH(1,1) model, the contribution to the overall kurtosis of the volatility clustering and of the innovations is symmetric.

The ACF of squared observations, derived by Taylor (1986), is given by

$$\rho_2(\tau) = \frac{\exp(\sigma_h^2 \phi^\tau) - 1}{\kappa_\varepsilon \exp(\sigma_h^2) - 1}, \quad \tau \geq 1. \tag{7}$$

If σ_h^2 is small and/or ϕ close to one, Taylor (1986) shows that the ACF in Equation (7) can be approximated by

$$\rho_2(\tau) \simeq \frac{\exp(\sigma_h^2) - 1}{\kappa_\varepsilon \exp(\sigma_h^2) - 1} \phi^\tau. \tag{8}$$
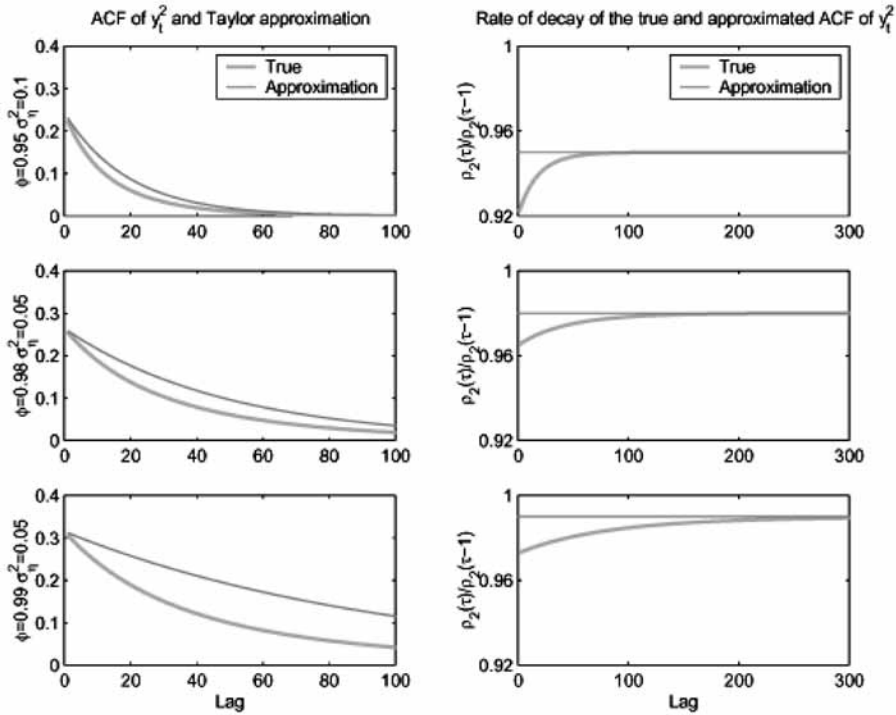
Figure 4 ACF of squared observations and rate of decay of such ACF for symmetric ARSV(1) models with gaussian innovations.

The pattern of the approximated autocorrelations in Equation (8) is the same as for the autocorrelations of an ARMA(1,1) model. Consequently, this approximation has often been used to argue that the ACF of squares in GARCH(1,1) and ARSV(1) models are similar. However, in general, the approximated autocorrelations are larger than the true ones and, for small lags, τ, their rate of decay is smaller than the rate of decay of the autocorrelations in Equation (7). As an illustration, Figure 4 plots the true and approximated ACF of squares together with their rates of decay for three alternative ARSV(1) models with parameters ($\phi = 0.95$, $\sigma_\eta^2 = 0.1$), ($\phi = 0.98$, $\sigma_\eta^2 = 0.05$), and ($\phi = 0.99$, $\sigma_\eta^2 = 0.05$), respectively. This figure shows that the use of the approximation could lead to a very distorting picture of the shape of the true ACF. However, since $\alpha + \beta$ and ϕ can be interpreted as measures of persistence in GARCH(1,1) and ARSV(1) models, respectively, many authors have compared their estimates [see, e.g., Taylor (1994) and Shephard (1996)]. Following them and given that, as Figure 4 illustrates, the rate of decay of the autocorrelations in Equation (7) tends for long lags to ϕ, we also consider this parameter as a measure of the persistence of shocks to volatility in ARSV(1) models [see also Kim, Shephard, and Chio (1998) and Meyer and Yu (2000)].
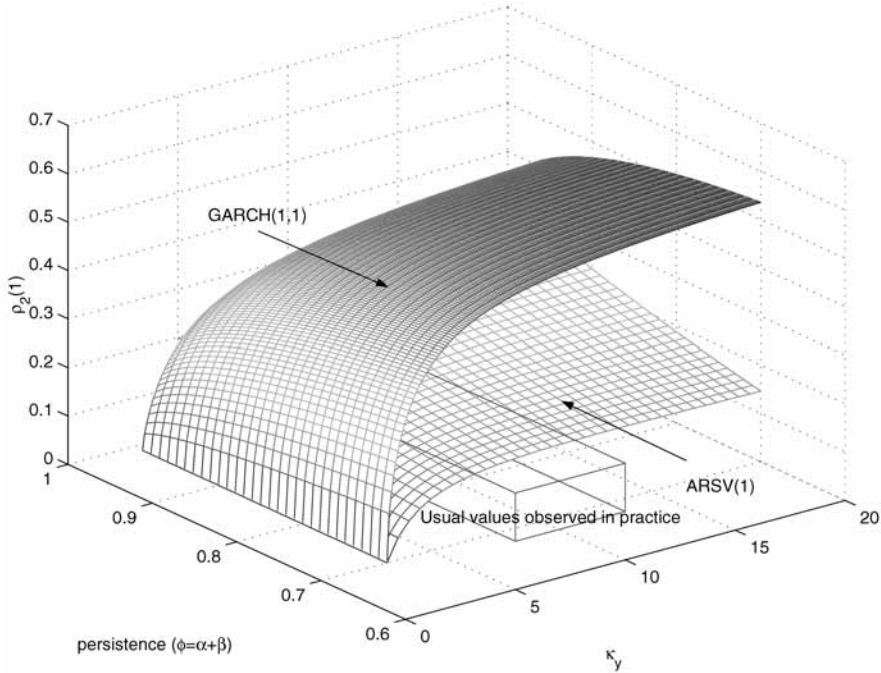
Figure 5 Relationship between kurtosis, first-order autocorrelation of squared observations, and persistence for GARCH(1,1) and ARSV(1) models with Gaussian errors.

The relationship between kurtosis, persistence, and $\rho_2(1)$ for an ARSV(1) model is given by

$$\rho_2(1) = \frac{\left(\frac{\kappa_y}{\kappa_\varepsilon}\right)^\phi - 1}{\kappa_y - 1}. \tag{9}$$

This relationship is plotted in Figure 5 for the normal-ARSV(1) model, together with the corresponding relationship for the normal-GARCH(1,1) model. As in Figure 3, the box plotted in Figure 5 contains the values of κ_y and $\rho_2(1)$ usually encountered in empirical applications. Several conclusions can be drawn from this figure. First, it is rather clear that ARSV(1) models are able to generate series with simultaneously higher kurtosis and lower $\rho_2(1)$ than the GARCH(1,1) model for a larger range of values of the persistence. Second, it is possible to observe that, for the same kurtosis and persistence, the autocorrelations of squared observations implied by the normal-ARSV(1) model are smaller than the autocorrelations of the normal-GARCH(1,1) model, except when the volatility approaches the nonstationary region, where both autocorrelations are the same. Furthermore, in ARSV(1) models, for a given kurtosis, the first-order autocorrelation of squares increases with the persistence parameter ϕ, while in GARCH(1,1) models the first-order autocorrelation decreases with $\alpha+\beta$. Consequently it could be expected that, for a given series of returns with a given kurtosis, if the first-order

autocorrelation of squares is small, the persistence estimated in ARSV(1) models is usually lower than in GARCH(1,1) models; observe the values of κ_y and $\rho_2(1)$ within the box usually observed in empirical applications. Therefore it is possible to have ARSV(1) models with high kurtosis, low $\rho_2(1)$, and persistence far from the nonstationary region, while in a normal-GARCH(1,1) model, the persistence would be high because it is the only way to achieve both high kurtosis and low $\rho_2(1)$.

Alternatively, as we mentioned before, the GARCH model can generate higher kurtosis without increasing the first-order autocorrelation of squares by having a conditional heavy-tailed distribution. For example, comparing the surfaces of the relationship between κ_y, $\rho_2(1)$ and persistence for the Student-GARCH(1,1) model with 7 degrees of freedom and the normal-ARSV(1) models, which appear in Figures 3 and 5, respectively, it is possible to observe that these two models are able to generate series with values of the kurtosis and $\rho_2(1)$ within the region defined by the box, where the typical values would be found. Therefore both models may have similar fits when implemented to represent the dynamic evolution of volatility of the same real-time series [see, e.g., Shephard (1996)]. However, remember that the restrictions on the α parameter of the Student-GARCH model, needed to guarantee that the fourth-order moment is finite, restrict severely the dynamic evolution of the volatility.

Summarizing, the ARSV(1) model does not force ε_t to have fat tails or ϕ to be close to one in order to allow simultaneously for high kurtosis and small $\rho_2(1)$. Therefore, introducing the noise η_t makes the ARSV(1) model more flexible even when compared with GARCH(1,1) models with conditional fat-tailed distributions.

1.3   Monte Carlo Simulation

In this section we illustrate with simulated data that when normal-GARCH models are fitted to represent the evolution of volatility, the kurtosis and first-order autocorrelation of squared returns implied by the estimated parameters could be much larger than the corresponding population coefficients of the simulated data. Furthermore, we also show that the persistence estimated by the GARCH model is usually greater than the persistence of the underlying true autocorrelations of squares. For this purpose, we generate 1000 replicates of size $T = 5000$ by different normal-ARSV(1) models and fit a stationary normal-GARCH(1,1) model to each of the simulated series. The parameters of the ARSV(1) models considered appear in Table 2 along with the corresponding population kurtosis and first-order autocorrelation of squares. These models have been chosen to generate simulated series with properties similar to those of real series of daily financial returns. Notice that, with the exception of the last model, the coefficient of variation defined as $\frac{\text{var}(\sigma_t^2)}{[E(\sigma_t^2)]^2} = \exp\left(\frac{\sigma_\eta^2}{1-\phi^2}\right) - 1$, always takes values between 0.33 and 2.53; for example, Jackier, Polson, and Rossi (1994) report several empirical studies where this coefficient is between 0.1 and 2. Finally, with the exception of the last two models, there are always two models with the same

Table 2 Means and standard deviations of estimates of the normal-GARCH(1,1) model when the series are generated by normal-ARSV(1) models.

| Simulated model ARSV(1) | | | | Estimated model GARCH(1,1) | | | | | | |
ϕ	σ_η^2	κ_y	$\rho_2(1)$	$\hat\omega$	$\hat\alpha$	$\hat\beta$	$\hat\alpha+\hat\beta$	%∄κ_y	κ_y	$\rho_2(1)$
0.8250	0.0919	4.0000	0.0893	0.2117 (0.0545)	0.0872 (0.0155)	0.7301 (0.0566)	0.8172 (0.0471)	0.0	3.1519 (0.0482)	0.1042 (0.0189)
0.9590	0.0231	4.0000	0.1059	0.0452 (0.0111)	0.0727 (0.0093)	0.8887 (0.0157)	0.9615 (0.0097)	0.0	3.5318 (0.1983)	0.1337 (0.0230)
0.9750	0.0252	5.0000	0.1614	0.0250 (0.0070)	0.0917 (0.0096)	0.8910 (0.0117)	0.9827 (0.0058)	5.90	9.9686 (32.741)	0.2697 (0.0503)
0.9000	0.1000	5.0781	0.1486	0.1210 (0.0255)	0.1378 (0.0168)	0.7730 (0.0302)	0.9108 (0.0202)	0.0	3.9523 (0.4045)	0.2183 (0.0313)
0.9165	0.1109	6.0000	0.1775	0.0991 (0.0198)	0.1653 (0.0177)	0.7722 (0.0244)	0.9376 (0.0153)	1.20	10.389 (114.94)	0.3108 (0.0455)
0.9798	0.0277	6.0000	0.1944	0.0188 (0.0058)	0.1023 (0.0098)	0.8876 (0.0109)	0.9898 (0.0046)	58.50	48.3644 (338.07)	0.3343 (0.0432)
0.9400	0.1401	7.0000	0.2334	0.0645 (0.0136)	0.2248 (0.0198)	0.7600 (0.0203)	0.9848 (0.0084)	99.30	81.3991 (74.8355)	0.4335 (0.0187)
0.9856	0.0344	7.0000	0.2529	0.0121 (0.0046)	0.1190 (0.0101)	0.8776 (0.0106)	0.9966 (0.0020)	100		
0.9500	0.1000	8.3666	0.2239	0.0513 (0.0116)	0.1889 (0.0169)	0.7940 (0.0179)	0.9829 (0.0082)	94.90	65.1018 (143.0182)	0.4024 (0.0255)
0.9000	0.2000	8.5955	0.2079	0.1247 (0.0252)	0.2393 (0.0252)	0.7089 (0.0286)	0.9483 (0.0179)	63.90	234.6440 (3470.1)	0.4227 (0.0386)
0.9800	0.0500	10.6041	0.2547	0.0173 (0.0056)	0.1418 (0.0116)	0.8543 (0.0121)	0.9960 (0.0022)	100		
0.9500	0.200	23.3336	0.2695	0.0561 (0.0130)	0.2650 (0.0203)	0.7319 (0.0205)	0.9969 (0.0015)	100		

kurtosis and different parameter ϕ, and therefore different rates of decay of the autocorrelations of squares toward zero. Table 2 reports the Monte Carlo means and standard deviations of the QML estimates of the GARCH(1,1) parameters α, β, and ω, as well as the estimated persistence $\alpha + \beta$. We also report the percentage of series where the estimated GARCH parameters do not satisfy the condition for the existence of the kurtosis. The mean of the implied kurtosis is computed based on the estimates that satisfy the condition. Finally, the mean of the implied $\rho_2(1)$ is shown in the last column.

Looking at the results reported in Table 2, it is possible to conclude that, as expected from the previous results, the persistence implied by the estimated GARCH(1,1) models is larger than the actual persistence of the series, especially for the series with higher kurtosis [see also Hafner and Herwartz (2000) for a similar result]. For example, in the ARSV(1) model with $\phi = 0.94$ and $\sigma_\eta^2 = 0.14$, the average estimated persistence is $\hat{\alpha} + \hat{\beta} = 0.985$ and in the model with $\phi = 0.98$ and $\sigma_\eta^2 = 0.05$, the Monte Carlo mean of $\hat{\alpha} + \hat{\beta}$ is 0.996.

Even more interesting is to observe the large number of series for which the corresponding GARCH(1,1) estimates do not satisfy the condition for the existence of kurtosis. In several cases the percentage is actually 100. This implies that, although the simulated series have perfectly defined finite kurtosis, looking at the estimated GARCH parameters, it seems that the fourth-order moment does not exist. For example, consider the ARSV model with parameters $\phi = 0.9856$ and $\sigma_\eta^2 = 0.0344$, which is highly plausible from an empirical point of view. In this case, although the kurtosis of the simulated series is seven, when the GARCH model is fitted to these series, the estimated parameters are such that it seems that the kurtosis is not finite for any of the series. This suggests that it is possible that the empirical conclusion often obtained, that the kurtosis of financial returns is not defined, could be a consequence of the rigidity of the GARCH model used to represent the dynamic evolution of volatility and not that the kurtosis of financial returns is not really defined [see, e.g., Jondeau and Rockinger (2003) and the references therein]. Furthermore, notice that, in most of the cases where the kurtosis exists, the kurtosis implied by the estimated GARCH parameters is clearly over the actual values. Consequently it seems clear that the GARCH model has serious difficulties representing the values of the kurtosis usually observed in daily financial returns. Finally, notice that when comparing models with the same kurtosis, the percentage of series where the estimated GARCH parameters do not satisfy the condition for the existence of the fourth-order moment is larger, the larger the parameter ϕ that measures the persistence.

Finally, Table 2 shows that the mean of the $\rho_2(1)$ coefficients implied by the GARCH estimates is always larger than the theoretical one.

2   PROPERTIES OF ASYMMETRIC MODELS

As we mentioned in the introduction, it is not unusual to find series of returns characterized by the asymmetric response of volatility to positive and negative

returns. In this section we analyze whether introducing this asymmetry changes the relationship between kurtosis, autocorrelation of squares, and persistence found in GARCH and ARSV models. There are a large number of asymmetric conditional heteroscedastic models proposed in the literature; see Henstchel (1995) and Duan (1997) for two models that embody many of the most popular specifications. In this section we consider the simplest specifications directly related with the GARCH(1,1) and ARSV(1) models analyzed before. In particular, we consider the QGARCH model proposed by Sentana (1995), the EGARCH model of Nelson (1991), and the asymmetric ARSV model proposed by Harvey and Shephard (1996).

2.1 QGARCH(1,1)

If y_t follows a QGARCH(1,1) model, then the volatility is given by

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma y_{t-1}, \tag{10}$$

where the positivity of σ_t^2 is guaranteed if ω, α, $\beta > 0$ and $\gamma^2 \leq 4\alpha\omega$. The process is covariance stationary if $p < 1$ [see He and Teräsvirta (1999)]. Notice that the covariance stationarity of y_t does not depend on the parameter γ that measures the asymmetry.

Sentana (1995) shows that the properties of the GARCH(1,1) and QGARCH(1,1) models are very similar. In fact, both have the same unconditional mean and variance equal to zero and $\sigma_y^2 = \frac{\omega}{1-p}$, respectively. Furthermore, in both models the odd moments are zero, the series y_t is uncorrelated, and the cross-correlations between y_t^2 and y_{t-h} are zero for all h except for $h=1$. In this case, $\text{cov}(y_t^2, y_{t-1}) = \gamma \sigma_y^2$ in the QGARCH(1,1) model and zero in the GARCH(1,1) model. Using the results of He and Teräsvirta (1999), it is possible to derive the following expression of the kurtosis of y_t,

$$\kappa_y = \kappa_\varepsilon \left[1 - \frac{\alpha^2(\kappa_\varepsilon - 1)}{1 - p^2} \right]^{-1} + \kappa_\varepsilon \frac{A^*}{1 - \alpha^2(\kappa_\varepsilon - 1) - p^2} \tag{11}$$

where $A^* = (\gamma/\sigma_y)^2$ and κ_ε is the kurtosis of ε_t. Notice that the kurtosis of y_t in the QGARCH models is always larger than in the corresponding symmetric GARCH model in Equation (3). However, this increase is small because, typically, in real-time series, γ is small relative to σ_y^2, and consequently A^* is rather small.

The ACF of y_t^2 is given by

$$\rho_2(\tau) = \begin{cases} \dfrac{2\alpha(1 - p^2 + \alpha p) + A^*(\kappa_\varepsilon \alpha + \beta)}{2(1 - p^2 + \alpha^2) + \kappa_\varepsilon A^*}, & \tau = 1 \\ p^{\tau-1}\rho_2(1), & \tau > 1 \end{cases} \tag{12}$$

This ACF decays at the same rate as in the GARCH(1,1) model. Therefore, once more the persistence can be measured by p. Furthermore, if, as mentioned before, A^* is rather small, the autocorrelation of order one is almost the same in both models.

Using Equations (11) and (12) and after some tedious algebra, it is possible to derive the following relationship between kurtosis, persistence, and $\rho_2(1)$ in a QGARCH(1,1) model:

$$\rho_2(1) = \frac{2\sqrt{\dfrac{(\kappa_y - \kappa_\varepsilon)(1-p^2) - \kappa_\varepsilon A^*}{(\kappa_\varepsilon - 1)\kappa_y}}\left[1 - p^2 + p\sqrt{\dfrac{(\kappa_y - \kappa_\varepsilon)(1-p^2) - \kappa_\varepsilon A^*}{(\kappa_\varepsilon - 1)\kappa_y}}\right]}{2\left[1 - p^2 + \dfrac{(\kappa_y - \kappa_\varepsilon)(1-p^2) - \kappa_\varepsilon A^*}{(\kappa_\varepsilon - 1)\kappa_y}\right] + \kappa_\varepsilon A^*}$$

$$+ \frac{A^*\left[p + (\kappa_\varepsilon - 1)\sqrt{\dfrac{(\kappa_y - \kappa_\varepsilon)(1-p^2) - \kappa_\varepsilon A^*}{(\kappa_\varepsilon - 1)\kappa_y}}\right]}{2\left[1 - p^2 + \dfrac{(\kappa_y - \kappa_\varepsilon)(1-p^2) - \kappa_\varepsilon A^*}{(\kappa_\varepsilon - 1)\kappa_y}\right] + \kappa_\varepsilon A^*}.$$

Figure 6 plots this relationship for a normal-QGARCH(1,1) model when $\alpha + \beta = 0.9$ and $A^* = 0$ and $A^* = 0.1$ and when $\alpha + \beta = 0.99$ for both values of A^*. As expected, this figure illustrates that, in usual cases in which the kurtosis takes values between 5 and 10, introducing asymmetry in a GARCH(1,1) model does not have any significant effect on the relationship between the three quantities of interest, apart from imposing more restrictions on the parameters
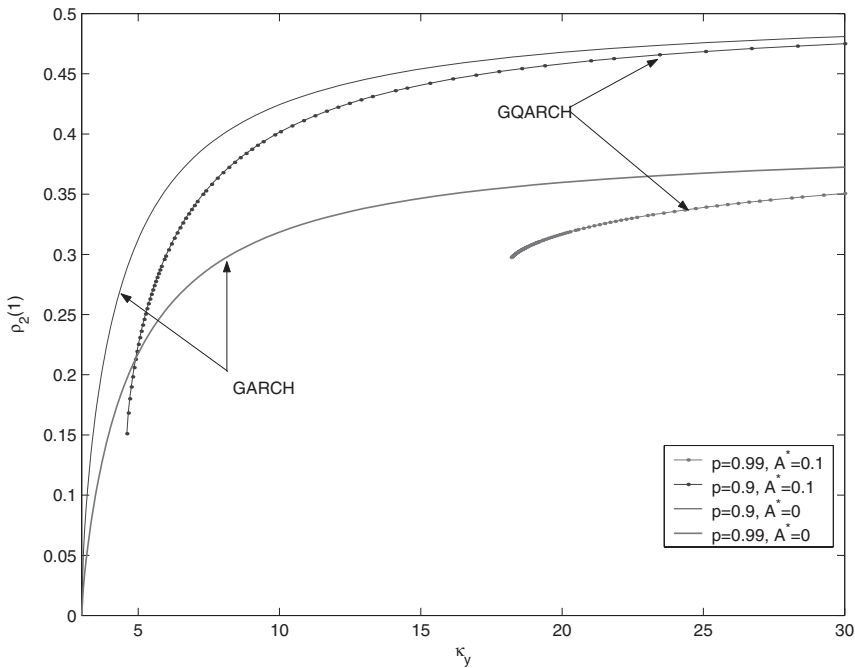

Figure 6 Relationship between kurtosis, first-order autocorrelation of squared observations, and persistence for asymmetric QGARCH(1,1) models.

in order to guarantee that the kurtosis is finite. For example, if $\alpha + \beta = 0.99$ and $A^* = 0.1$, the kurtosis implied by the QGARCH model is always greater than 17.

2.2 EGARCH(1,0)

The EGARCH model was proposed by Nelson (1991) to overcome some of the limitations of the GARCH models. In particular, he points out that the constraints to guarantee the positivity of the conditional variance restrict the dynamics of the volatility and are not always satisfied in empirical applications. Furthermore, the definition of persistence could be rather confusing because shocks to volatility may persist in one norm and die out in another. Finally, the GARCH(1,1) model does not represent the asymmetric response of volatility. The specification of the volatility in the EGARCH model can be considered in between the GARCH and ARSV specifications in the sense that it has a unique disturbance, as in the GARCH model, but the logarithmic transformation of the volatility is modeled, as in the ARSV model.

The volatility of the simplest EGARCH(1,0) model is given by

$$\log \sigma_t^2 = \omega + \beta \log \sigma_{t-1}^2 + \alpha[|\varepsilon_{t-1}| - E(|\varepsilon_{t-1}|)] + \gamma \varepsilon_{t-1}, \tag{14}$$

where γ is the parameter that measures the asymmetry. The model is stationary if $|\beta| < 1$. The unconditional variance, kurtosis, and autocorrelations of squares can be derived using the results in Nelson (1991) and are given, respectively, by

$$\sigma_y^2 = \exp\left(\frac{\omega}{1-\beta}\right) \prod_{i=1}^{\infty} E(\exp(\beta^{i-1}g)), \tag{15}$$

$$\kappa_y = \kappa_\varepsilon \prod_{i=1}^{\infty} \frac{E(\exp(2\beta^{i-1}g))}{[E(\exp(\beta^{i-1}g))]^2}, \tag{16}$$

and

$$\rho_2(\tau) = \frac{E(\varepsilon_{t-1}^2 \exp(\beta^{\tau-1}g))P_1 P_2 - P_3}{\kappa_\varepsilon P_4 - P_3}, \tag{17}$$

where $g = \alpha[|\varepsilon_{t-1}| - E(|\varepsilon_{t-1}|)] + \gamma \varepsilon_{t-1}$, $P_1 = \prod_{i=1}^{\tau} E(\exp(\beta^{i-1}g))$, $P_2 = \prod_{i=1}^{\infty} E(\exp((1+\beta^\tau)\beta^{i-1}g))$, $P_3 = \prod_{i=1}^{\infty} [E(\exp(2\beta^{i-1}g))]^2$, and $P_4 = \prod_{i=1}^{\infty} E(\exp(2\beta^{i-1}g))$. Nelson (1991) has evaluated the expectations involved in Equations (15)–(17) for different distributions of ε_t. In particular, if ε_t is gaussian, then

$$E[\exp\{bg\}] = \{\Phi(bc_1)\exp\{0.5b^2c_1^2\} + \Phi(bc_2)\exp\{0.5b^2c_2^2\}\}\exp\{-b\alpha(2/\pi)^{1/2}\}$$

$$E[\varepsilon_{t-1}^2 \exp(bg)] = 2\sqrt{\frac{2}{\pi}}\exp\left\{-\sqrt{\frac{2}{\pi}}\alpha b\right\}\exp\{0.25c_1^2 b^2\}D_{(-3)}(-bc_1)$$

$$+ \exp\{0.25c_2^2 b^2\}D_{(-3)}(-bc_2),$$

where $c_1 = \alpha + \gamma$, $c_2 = \alpha - \gamma$, $D_{(q)}$ is the parabolic cylinder function and $\Phi(\bullet)$ is the standard normal cumulative distribution function. He Teräsvirta, and Malmsten (2002) compute the moments of EGARCH models in more general settings.
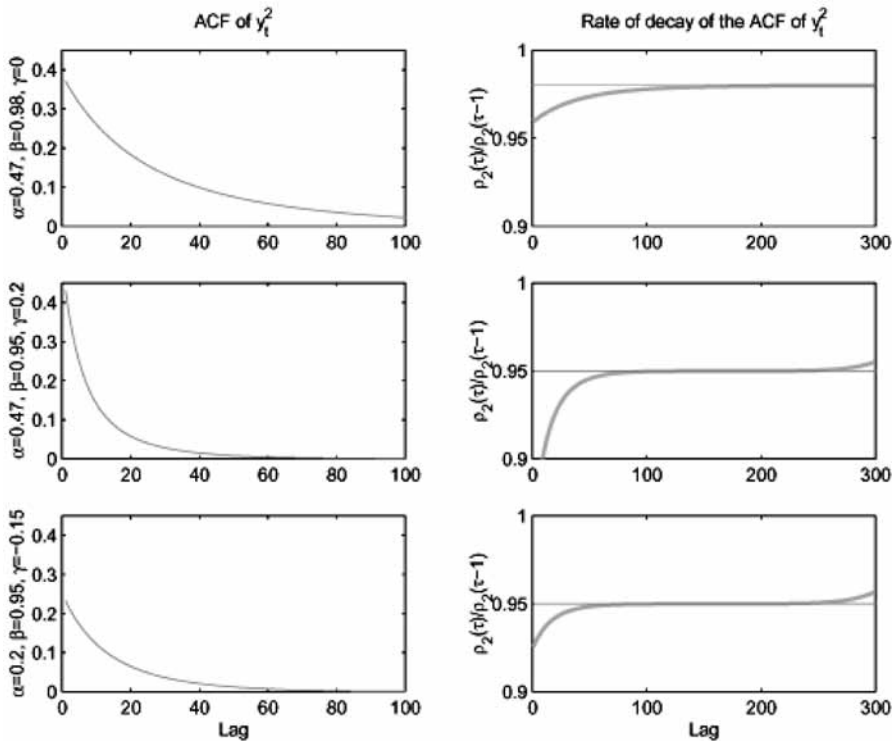
Figure 7 ACF of squared observations and rate of decay for EGARCH(1,0) models with gaussian innovations.

As an illustration, Figure 7 represents the ACF in Equation (17) together with its rate of decay for three selected EGARCH models with parameters $\{\alpha = 0.47, \beta = 0.98, \gamma = 0\}$, $\{\alpha = 0.47, \beta = 0.95, \gamma = 0.2\}$, and $\{\alpha = 0.2, \beta = 0.95, \gamma = -0.15\}$, respectively. It is interesting to observe that the rate of decay of the autocorrelations is not β, but depends on the other parameters of the model, α and γ. Finally, notice the similarity in the patterns of the rates of decay of the autocorrelations of squares in ARSV and EGARCH models; compare Figures 4 and 7.

Given that it is rather difficult to find the expression that relates the kurtosis, persistence, and autocorrelation of order one of squares in an EGARCH(1,0) model, Figure 8 plots the function that relates the kurtosis and $\rho_2(1)$ for different parameter values. This figure shows that this relationship is rather similar in GARCH and EGARCH models. For a given kurtosis, the $\rho_2(1)$ coefficient is smaller, the larger is the persistence parameter, β. Although the introduction of the asymmetry parameter allows us to generate series with a larger range of possible values of the kurtosis and autocorrelations of squares, the values of $\rho_2(1)$ corresponding to values of the kurtosis usually observed in real-time series are even higher than in the corresponding GARCH models.
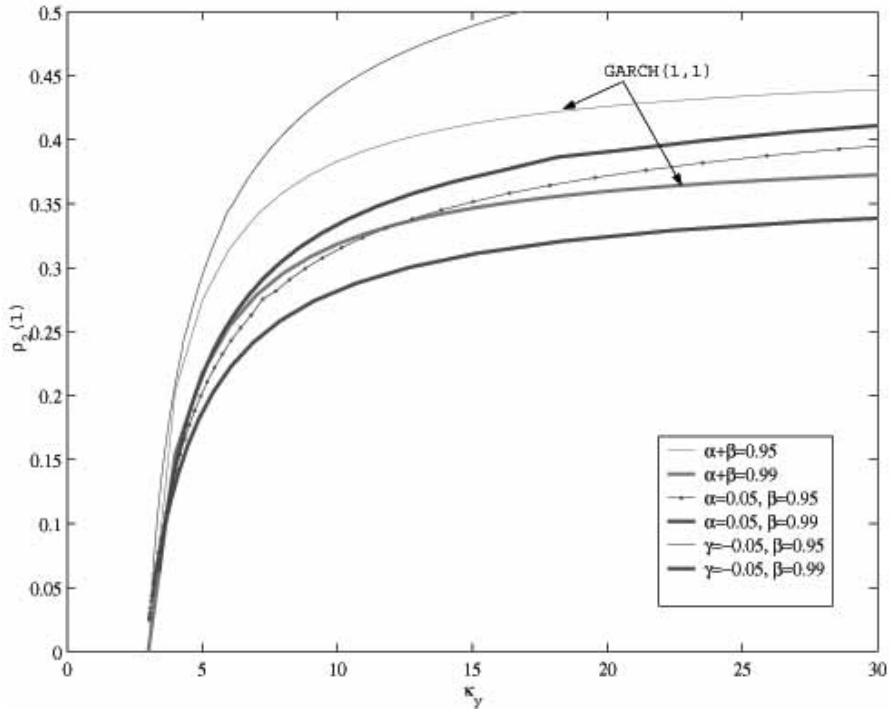
Figure 8 Relationship between kurtosis and first-order autocorrelation of squared observations for different values of the persistence for asymmetric EGARCH(1,0) models.

2.3 Asymmetric ARSV(1)

As far as we know, the only asymmetric ARSV model proposed in the literature at the moment is from Harvey and Shephard (1996), who introduce the asymmetry through correlation between the noises ε_t and η_t. Therefore the volatility of the asymmetric ARSV(1), A-ARSV(1), model is given by Equation (6) with corr$(\varepsilon_t, \eta_t) = \delta$. Harvey and Shephard (1996) show that, in this case, the kurtosis of y_t is the same as in the symmetric case. The ACF of squared observations, derived by Taylor (1994), is given by

$$\rho_2(\tau) = \frac{(1 + \delta^2 \sigma_\eta^2) \exp(\sigma_h^2 \phi^\tau) - 1}{\kappa_\varepsilon \exp(\sigma_h^2) - 1}, \quad \tau \geq 1. \tag{18}$$

As an illustration, Figure 9 plots the ACF of squares and its rate of decay for three alternative A-ARSV(1) models with parameters $(\phi = 0.95, \sigma_\eta^2 = 0.1)$, $(\phi = 0.98, \sigma_\eta^2 = 0.05)$, and $(\phi = 0.99, \sigma_\eta^2 = 0.05)$, respectively, and for $\delta = 0.1$ and 0.9. It is interesting to observe that, as in the symmetric ARSV model, the rate of decay is less than ϕ for the smaller lags, but instead of converging to ϕ, it converges
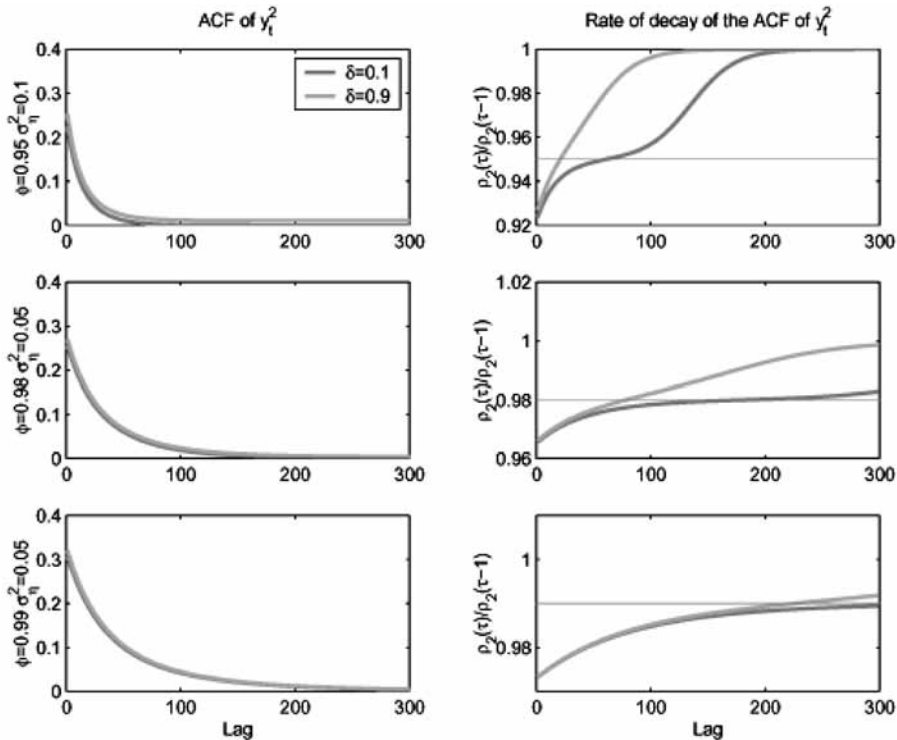
Figure 9 ACF of squared observations and rate of decay for asymmetric ARSV(1) models.

to one in the presence of correlation between the level and volatility noises, ε_t and η_t, respectively. However, in practice, notice that the autocorrelations of large order are indistinguishable from zero.

From Equation (18), it is straightforward to find $\rho_2(1)$ as a function of ϕ and κ_y, as follows

$$\rho_2(1) = \frac{\left[1 + \delta^2(1-\phi^2)\log\left(\frac{\kappa_y}{\kappa_\varepsilon}\right)\right]\left(\frac{\kappa_y}{\kappa_\varepsilon}\right)^\phi - 1}{\kappa_y - 1}. \tag{19}$$

Figure 10 represents this relationship for normal-A-ARSV(1) models for two different values of the asymmetry parameter, $\delta = 0.2$ and $\delta = 0.9$. It is possible to observe that, for a given value of the kurtosis, the autocorrelation of order one of squares, $\rho_2(1)$, is larger, the larger the correlation between the noises.

Therefore, if as expected in empirical applications, the magnitude of the asymmetry parameter is rather small, the relationship between persistence, kurtosis, and autocorrelations of squares is similar to the one derived for the symmetric ARSV(1) model. Introducing the asymmetric response of volatility to positive and negative returns does not change the main conclusions obtained in the previous section.
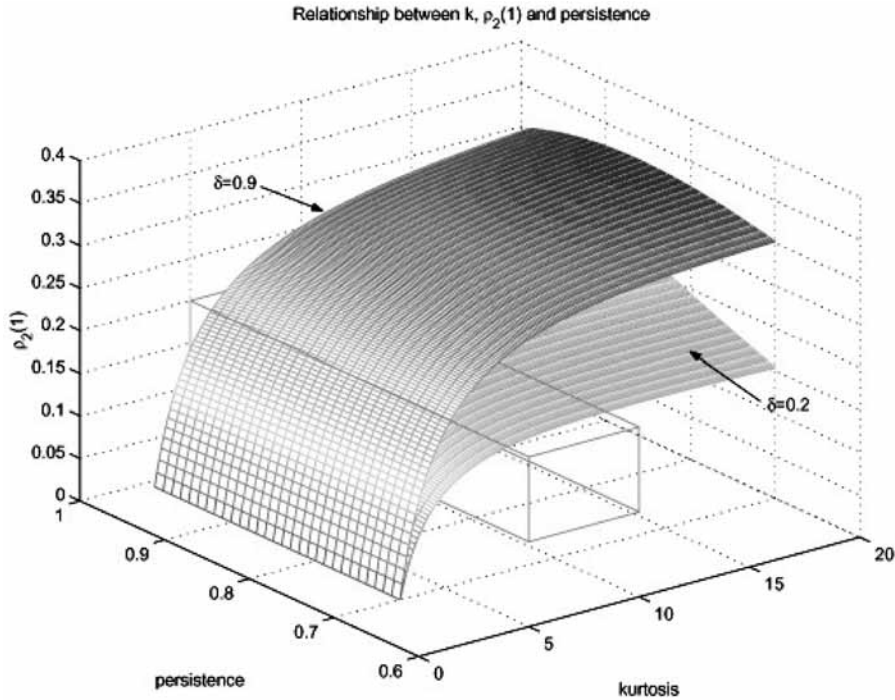
Figure 10 Relationship between kurtosis, first-order autocorrelation of squared observations, and persistence for asymmetric ARSV(1) models.

3   EMPIRICAL APPLICATION

In this section we illustrate the previous results by fitting GARCH(1,1), EGARCH(0,1), and ARSV(1) models to the four series of daily financial returns described in the introduction. The parameters of the EGARCH model have been estimated by maximum likelihood (ML) assuming conditional normality of the innovations, while those of the GARCH(1,1) model have been estimated assuming both a conditional normal and a Student's t distribution.[2] Finally, the ARSV model has been estimated using the method proposed by Sandman and Koopman (1998).[3] For each of the series, the kurtosis, persistence, and auto-correlations of squares implied by the corresponding estimated parameters are compared.

Table 3 reports the ML estimates of the parameters together with their asymptotic standard errors. Looking, for example, at the results for the US-JAP series, we conclude that the parameters that allow the volatility to evolve over time, α in the GARCH-type model and σ_η^2 in the ARSV model, are always

---

[2]   The estimates of the GARCH and EGARCH models were obtained using the Eviews 3.1 program.

[3]   The program used to estimate the ARSV parameters is $sv-mcl.ox$ and it is available at http://www.econ.vu.nl/koopman/sv/.

Table 3 Estimated models for the daily returns.

	GARCH	GARCH-t	EGARCH	ARSV
US-CAN				
	$\hat{\omega} = 0.0011$	$\hat{\omega} = 0.0010$	$\hat{\omega} = -0.1040$	$\hat{\sigma}_*^2 = 0.0775$
	(0.001)	(0.001)	(0.012)	(0.009)
	$\hat{\alpha} = 0.0442$	$\hat{\alpha} = 0.0497$	$\hat{\alpha} = 0.1010$	$\hat{\sigma}_\eta^2 = 0.0219$
	(0.005)	(0.009)	(0.011)	(0.008)
	$\hat{\beta} = 0.9449$	$\hat{\beta} = 0.9418$	$\hat{\beta} = 0.9890$	$\hat{\phi} = 0.9718$
	(0.006)	(0.011)	(0.003)	(0.011)
		$\hat{\nu} = 6.1275$	$\hat{\gamma} = 0.0140$	
			(0.008)	
Implied persistence	$\hat{\alpha} + \hat{\beta} = 0.9891$	$\hat{\alpha} + \hat{\beta} = 0.9915$	$\hat{\beta} = 0.9890$	$\hat{\phi} = 0.9718$
Implied κ_y	3.6595	19.6207	3.6190	4.4480
Implied $\rho_2(1)$	0.1223	0.1696	0.1176	0.1352
US-JAP				
	$\hat{\omega} = 0.0130$	$\hat{\omega} = 0.0057$	$\hat{\omega} = -0.1430$	$\hat{\sigma}_*^2 = 0.4217$
	(0.002)	(0.003)	(0.012)	(0.039)
	$\hat{\alpha} = 0.0559$	$\hat{\alpha} = 0.0400$	$\hat{\alpha} = 0.1620$	$\hat{\sigma}_\eta^2 = 0.1024$
	(0.006)	(0.009)	(0.013)	(0.036)
	$\hat{\beta} = 0.9230$	$\hat{\beta} = 0.9498$	$\hat{\beta} = 0.9610$	$\hat{\phi} = 0.9165$
	(0.007)	(0.012)	(0.005)	(0.027)
		$\hat{\nu} = 4.7460$	$\hat{\gamma} = -0.0400$	
			(0.007)	
Implied persistence	$\hat{\alpha} + \hat{\beta} = 0.9789$	$\hat{\alpha} + \hat{\beta} = 0.9898$	$\hat{\beta} = 0.9610$	$\hat{\phi} = 0.9165$
Implied κ_y	3.5281	53.0184	3.5237	5.6888
Implied $\rho_2(1)$	0.1202	0.1094	0.1344	0.1701
BOMBAY				
	$\hat{\omega} = 0.0702$	$\hat{\omega} = 0.2663$	$\hat{\omega} = -0.0320$	$\hat{\sigma}_*^2 = 2.1302$
	(0.011)	(0.106)	(0.011)	(0.226)
	$\hat{\alpha} = 0.0434$	$\hat{\alpha} = 0.1317$	$\hat{\alpha} = 0.1500$	$\hat{\sigma}_\eta^2 = 0.2513$
	(0.006)	(0.034)	(0.016)	(0.075)
	$\hat{\beta} = 0.9369$	$\hat{\beta} = 0.7993$	$\hat{\beta} = 0.9360$	$\hat{\phi} = 0.8499$
	(0.006)	(0.054)	(0.009)	(0.041)
		$\hat{\nu} = 4.0764$	$\hat{\gamma} = -0.0510$	
			(0.011)	
Implied persistence	$\hat{\alpha} + \hat{\beta} = 0.9803$	$\hat{\alpha} + \hat{\beta} = 0.9386$	$\hat{\beta} = 0.9360$	$\hat{\phi} = 0.8499$
Implied κ_y	3.3207		3.3061	7.4160
Implied $\rho_2(1)$	0.0866	0.2238	0.1064	0.1805
S & P 500				
	$\hat{\omega} = 0.0042$	$\hat{\omega} = 0.0026$	$\hat{\omega} = -0.0670$	$\hat{\sigma}_*^2 = 0.6001$
	(0.001)	(0.001)	(0.007)	(0.087)
	$\hat{\alpha} = 0.0326$	$\hat{\alpha} = 0.0360$	$\hat{\alpha} = 0.0860$	$\hat{\sigma}_\eta^2 = 0.0276$
	(0.003)	(0.002)	(0.009)	(0.008)

Table 3 (*continued*)

	GARCH	GARCH-t	EGARCH	ARSV
	$\hat{\beta} = 0.9619$	$\hat{\beta} = 0.9637$	$\hat{\beta} = 0.9890$	$\hat{\phi} = 0.9781$
	(0.003)	(0.002)	(0.001)	(0.007)
		$\hat{\nu} = 4.4786$	$\hat{\gamma} = -0.0460$	
			(0.006)	
Implied persistence	$\hat{\alpha} + \hat{\beta} = 0.9945$	$\hat{\alpha} + \hat{\beta} = 0.9997$	$\hat{\beta} = 0.9$	$\hat{\phi} = 0.9781$
Implied κ_y	3.7210		3.8644	5.6730
Implied $\rho_2(1)$	0.1176	0.6948	0.1335	0.1851

significant. Furthermore, the asymmetry parameter in the EGARCH model is also significant. The estimate of the persistence is smaller when the ARSV is fitted (0.92) than with the GARCH-type model which is closer to 1. On the other hand, the kurtosis implied by the normal-GARCH and EGARCH models are very similar (around 3.52), while that implied by the ARSV model (5.69) is closer to the sample kurtosis of 5.60. However, notice that the kurtosis implied by the Student-GARCH model, 53.02, is obviously too high to represent the empirical kurtosis. Finally, the implied first-order autocorrelation of squares is higher in the ARSV model than in the GARCH-type model. Notice that the properties implied by the EGARCH estimates are clearly closer to the properties of the GARCH model than to the properties of the ARSV models. The results and conclusions for the other three series are similar.

A very common diagnostic for heteroscedastic time-series models is based on the properties of the standardized observations. For the US-JAP series, the sample kurtosis of the observations standardized using the ARSV estimated volatilities is 3.19, while if the observations are standardized by the normal-GARCH or the EGARCH estimates of σ_t, their kurtosis is approximately 6. Therefore it seems that the normal-GARCH(1,1) and EGARCH models are not adequate to represent the values of the kurtosis often observed in real data, while the assumption of normality in the ARSV model seems more adequate. Finally, the kurtosis of the standardized observations in the Student-GARCH model is 6.21, which is in agreement with a Student's t distribution with 5 degrees of freedom.

4   CONCLUSION

In this article we show that the relationship between the persistence of shocks to volatility, the first-order autocorrelation of squared observations, and the kurtosis implied by GARCH and ARSV models is rather different. In order to have simultaneously high kurtosis and low first-order autocorrelation of squares, the GARCH model is forced to have a persistence close to the nonstationary region and/or leptokurtic conditional distributions. Alternatively, the range of values of the parameters of the ARSV model that can generate series with small autocorrelations and high kurtosis is larger than for the GARCH model. On top of that, the

restrictions to guarantee the existence of the fourth-order moment in GARCH models severely restrict the allowed dynamic dependence of the volatility, while they are not as severe in ARSV models. Finally, we show that incorporating the asymmetric response of volatility to positive and negative shocks does not change the main conclusions obtained for the symmetric models.

REFERENCES

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys. (2003). "Modeling and Forecasting Realized Volatility." *Econometrica 71*, 579–625.

Anderson, J. (2001). "On the Normal Inverse Gaussian Stochastic Volatility Model." *Journal of Business and Economic Statistics 19*, 44–54.

Bai, X., J. R. Russell, and G. Tiao. (2003). "Kurtosis of GARCH and Stochastic Volatility Models with Non-Normal Innovations." *Journal of Econometrics 114*, 349–360.

Bera, A. K., and M. L. Higgins. (1995). "On ARCH Models: Properties, Estimation and Testing." In L. Oxley, D. A. R. George, C. J. Roberts, and S. Sayer (eds.), *Surveys in Econometrics*. Oxford: Blackwell.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics 31*, 307–327.

Bollerslev, T. (1988). "On the Correlation Structure for the Generalized Autoregressive Conditional Heteroskedastic Process." *Journal of Time Series Analysis 9*, 121–131.

Bollerslev, T., R. F. Engle, and D. B. Nelson. (1994). "ARCH Models." In *Handbook of Econometrics*, Vol. 4.

Campbell, J. Y., and L. Hentchel. (1992). "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns." *Journal of Financial Economics 31*, 281–318.

Diebold, F. X., and J. A. López. (1995). "ARCH Models." In K. Hoover (eds.), *Macroeconomics: Developments, Tensions and Prospects*. Oxford: Blackwell.

Duan, J. C. (1997). "Augmented GARCH(p, q) Process and its Diffusion Limit." *Journal of Econometrics 79*, 97–127.

Engle, R. F. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica 50*, 987–1007.

Engle, R. F. (1994). *ARCH Models*. Oxford: Oxford University Press.

Ghysels, E., A. C. Harvey, and E. Renault. (1996). "Stochastic Volatility." In G. S. Maddala and C. R. Rao (eds.), *Handbook of Statistics*, Vol. 14. Amsterdam: North-Holland.

Hafner, C. M., and H. Herwartz. (2000). "Testing for Linear Autoregressive Dynamics under Heteroscedasticity." *Econometrics Journal 3*, 177–197.

Hall, P., and Q. Yao. (2003). "Inference in ARCH and GARCH Models with Heavy-Tailed Errors." *Econometrica 71*, 285–317.

Harvey, A. C., and N. G. Shephard. (1996). "Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns." *Journal of Business and Economic Statistics 14*, 429–434.

Harvey, A. C., and M. Streibel. (1998). "Testing for a Slowly Changing Level with Special Reference to Stochastic Volatility." *Journal of Econometrics 87*, 167–189.

He, C., and T. Teräsvirta. (1999). ''Properties of Moments of a Family of GARCH Processes.'' *Journal of Econometrics 91*, 173–192.

He, C., T. Teräsvirta, and H. Malmsten. (2002). ''Moment Structure of a Family of First-Order Exponential GARCH Models.'' *Econometric Theory 18*, 868–885.

Hentschel, L. (1995). ''All in the Family: Nesting Symmetric and Asymmetric GARCH Models.'' *Journal of Financial Economics 39*, 71–104.

Jackier, E., N. G. Polson, and P. E. Rossi. (1994). ''Bayesian Analysis of Stochastic Volatility Models (with discussion).'' *Journal of Business and Economic Statistics 12*, 371–417.

Jondeau, E., and M. Rockinger. (2003). ''Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements.'' *Journal of Economic Dynamics & Control 27*, 1699–1737.

Kim, S., N. Shephard, and S. Chib. (1998). ''Stochastic Volatility: Likelihood Inference and Comparison with ARCH Model.'' *Review of Economic Studies 65*, 361–393.

Liesenfeld, R., and R. C. Jung. (2000). ''Stochastic Volatility Models: Conditional Normality Versus Heavy-Tailed Distributions.'' *Journal of Applied Econometrics 15*, 137–160.

Loudon, G. F., W. H. Watt, and P. K. Yadav. (2000). ''An Empirical Analysis of Alternative Parametric ARCH Models.'' *Journal of Applied Econometrics 15*, 117–136.

McAleer, M., and L. Oxley. (2003). *Contributions to Financial Econometrics*. Oxford: Blackwell.

Meyer, R., and J. Yu. (2000). ''BUGS for a Bayesian Analysis of Stochastic Volatility Models.'' *Econometrics Journal 3*, 198–215.

Nelson, D.B. (1991). ''Conditional Heteroskedasticity in Asset Returns: A New Approach.'' *Econometrica 59*, 347–370.

Sandmann, G., and S. J. Koopman. (1998). ''Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood.'' *Journal of Econometrics 87*, 271–301.

Sentana, E. (1995). ''Quadratic ARCH models.'' *Review of Economic Studies 62*, 639–661.

Shephard, N. G. (1996). ''Statistical Aspects of ARCH and Stochastic Volatility.'' In D. R. Cox, D. V. Hinkley, and O. E. Bardorff-Nielsen (eds.). London: Chapman and Hall.

Taylor, S. J. (1986). *Modelling Financial Time Series*. New York: John Wiley & Sons.

Taylor, S. J. (1994). ''Modelling Stochastic Volatility: A Review and Comparative Study.'' *Mathematical Finance 4*, 183–204.

Teräsvirta, T. (1996). ''Two Stylized Facts and the GARCH(1,1) Model.'' Working paper, Stockholm School of Economics.