

SIGN- AND VOLATILITY-SWITCHING ARCH MODELS: THEORY AND APPLICATIONS TO INTERNATIONAL STOCK MARKETS

FABIO FORNARI^{a*} AND ANTONIO MELE^b

^a*Banca d'Italia, Research Department, Via Nazionale 91, 00184 Roma, Italy*

^b*University of Paris X, MODEM, Avenue de la République 200, 92001 Nanterre, France*

SUMMARY

This paper develops two conditionally heteroscedastic models which allow an asymmetric reaction of the conditional volatility to the arrival of news. Such a reaction is induced by both the sign of past shocks and the size of past unexpected volatility. The proposed models are shown to converge in distribution to absolutely continuous Itô diffusion processes, as happens for other heteroscedastic formulations. One of the schemes developed in the paper—the Volatility-switching ARCH—differs from the existing asymmetric models insofar as it is able to capture a particular aspect of the behaviour of the volatilities, i.e. the reversion of their asymmetric reaction to news. Empirical evidence from stock market returns in six countries shows that such a model outperforms traditional asymmetric ARCH equations. © 1997 by John Wiley & Sons, Ltd. *J. appl. econom.* 12: 49–65, 1997.

(No. of Figures: 1. No. of Tables: 4. No. of Refs: 40.)

1. INTRODUCTION

Since Engle's (1982) and Bollerslev's (1986) seminal papers, ARCH (AutoRegressive Conditional Heteroscedastic) models have been widely employed in the analyses of financial markets. The effects of heteroscedasticity have been evidenced especially for high-frequency returns, whose distributions are heavy-peaked and tailed.¹

The original ARCH model posits the existence of a relation between past squared innovations of an observation assets returns changes model and their current conditional variances. Let ε_t be the innovation of an observation model; then, the GARCH(1,1) model assumes that ε_t is conditionally normal with variance changing through time in a fashion which resembles a restricted ARMA process, i.e.

$$\varepsilon_t | I_{t-1} \sim N(0, \sigma_t^2) \quad (1)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (2)$$

where $\alpha_0 > 0$, $\alpha_1, \beta \geq 0$ are real, non-stochastic parameters and I_{t-1} is the information set dated $t - 1$.

A shortcoming of the GARCH model is that the sign of the forecast errors does not influence the conditional variance, which may contradict the observed dynamics of assets returns. Black (1976), for example, noted that volatility tends to grow in reaction to bad news (excess returns

* Correspondence to: F. Fornari, Banca d'Italia Research Department CFI, Via Nazionale 91, D0184 Roma, Italy.

¹ For an extensive review of ARCH models, see Bera and Higgins (1993).

lower than expected), and to fall in response to good news (excess returns higher than expected). The economic explanation given by Black is that negative (positive) excess returns make the equity value, hence the leverage ratio, of a given firm increase (fall), thus raising (lowering) its riskiness and the future volatility of its assets. This phenomenon has consequently come to be referred to as the *leverage effect* (Pagan and Schwert, 1990; Campbell and Hentschel, 1992).

The basic attempts to include such features of the returns of financial assets into a convenient econometric framework are the Exponential ARCH model of Nelson (1991), the Threshold ARCH model of Zakoian (1994) and Rabemananjara and Zakoian (1993), the Asymmetric Power ARCH model of Ding *et al.* (1993), and the Stochastic Variance model of Harvey *et al.* (1994), or Harvey and Shephard (1993a,b). All such models include the sign of past forecast errors as conditioning information for the current values of the conditional variance.

The main concern of this paper is to develop heteroscedastic formulations which turn out to be useful in modelling the statistical properties of financial data. It improves over previously developed models for two main reasons:

- First, it develops a class of asymmetric ARCH models in which volatility is influenced by the sign of previous shocks and the unexpected volatility induced by such shocks.
- Second, it derives the asymptotic properties of such models, useful in the estimation of continuous time models recently developed in finance (see e.g. Hull and White, 1987; Longstaff and Schwartz, 1992; Fornari and Mele, 1995b).

With concern for the first issue, we propose two new models. In the first the intercept of the volatility equation — α_0 in equation (2) — is allowed to change according to the sign of previous shocks, so capturing the asymmetry of the conditional volatility within a simple traditional GARCH structure. Since the model is closely related to the Sign Conditional Autoregressive Model reported in Granger and Teräsvirta (1993), it will be referred to as Sign-switching ARCH. However, there are reasons to believe that factors other than the sign of past shocks are responsible for the asymmetric behaviour of volatilities. To examine such an opportunity, we propose a model — the Volatility-switching ARCH — which captures asymmetries via the impact of past shocks on the level of the volatility, rather than through the unexpected returns. Unlike previous models, it is able to capture an already observed phenomenon, the reversal of asymmetry, which will be defined in the next paragraph.

Second, it has been widely recognized that many of the GARCH models developed so far admit a continuous time representation, thus being useful also to estimate continuous time models employed in finance. Following this stream of research, analogous results are presented for the GARCH models hereby developed, as well as for other discrete time ARCH models.²

The paper is structured as follows. The next section deals with the Sign-switching and the Volatility-switching ARCH models and derives the expressions of their first four moments. We conclude the section by presenting weak convergence results for the Sign-switching ARCH, the Volatility-switching model and the Glosten *et al.* (1993; henceforth GJR) model. Section 3 presents our empirical results. Evidence from six daily stock market indices unambiguously shows that the Sign- and Volatility-switching ARCH models successfully detect asymmetries — and reversals — in the time series of the conditional volatility. Section 4 presents conclusions.

² This kind of result was pioneered by Nelson (1990), who showed that the GARCH(1,1) model of Bollerslev (1986) and the AR(1)-Exponential ARCH model of Nelson (1991) approach continuous-time AR(1) processes, as the length of the sample frequency approaches zero. Later, Fornari and Mele (1995a) extended his results to the case of the Asymmetric Power ARCH model of Ding *et al.* (1993), and El Babsiri and Zakoian (1994) established weak convergence theory for the Threshold ARCH model of Zakoian (1994).

2. SIGN- AND VOLATILITY-SWITCHING ARCH MODELS

2.1. The Structure of the Models

In the Sign-switching ARCH model we capture the asymmetric reaction of the conditional variance to shocks of different sign through the sign of such shocks. Let ε_t be a (scalar) innovation of a given (unidimensional) observation model. The Sign-switching GARCH(p, q, y) model then assumes:

$$\varepsilon_t \equiv z_t \sigma_t, \quad \varepsilon_t | I_{t-1} \sim N(0, \sigma_t^2) \quad (3)$$

$$\sigma_t^2 = w + \sum_{i=1, \dots, p} \beta_i \sigma_{t-i}^2 + \sum_{j=1, \dots, q} \alpha_j \varepsilon_{t-j}^2 + \sum_{x=1, \dots, y} \Phi_x s_{t-x} \quad (4)$$

$$s_t = +1 \quad \text{if } \varepsilon_t > 0 \quad (5a)$$

$$s_t = 0 \quad \text{if } \varepsilon_t = 0 \quad (5b)$$

$$s_t = -1 \quad \text{if } \varepsilon_t < 0 \quad (5c)$$

where p, q and $y \geq 0$, w, α_j ($j = 1, \dots, q$), β_i ($i = 1, \dots, p$) and Φ_x ($x = 1, \dots, y$) are real, non-stochastic parameters, satisfying $w > 0$, $\alpha_j \geq 0$, $\beta_i \geq 0$ and, finally, $|\sum_x \Phi_x \leq w|$; such constraints guarantee that the process $\{\sigma_t^2\}$ almost certainly remains positive.

Throughout the paper we will confine ourselves to the special case $p = q = y = 1$, so that equation (4) reduces to:

$$\sigma_t^2 = w + \beta \sigma_{t-1}^2 + \alpha \varepsilon_{t-1}^2 + \Phi s_{t-1} \quad (6)$$

It is straightforward to see that according to equations (3)–(6) one captures asymmetric responses of the volatility to positive and negative shocks³ since, when $\Phi < 0$, negative (positive) shocks observed at $t - 1$ will be associated with a higher (lower) level of the volatility at t .

The second and fourth unconditional moments of the innovations of the Sign-switching(1,1,1) model are (see Appendix 1):

$$E(\varepsilon^2) = w(1 - \alpha - \beta)^{-1} \quad (7)$$

$$E(\varepsilon^4) = 3 \frac{(w^2 + \Phi^2)(1 - \alpha - \beta) + 6w^2(\alpha + \beta)}{(1 - \alpha - \beta)(1 - \alpha^2 - 3\beta^2 - 2\alpha\beta)} \quad (8)$$

While the second unconditional moment coincides with that of a GARCH(1,1) (Bollerslev, 1986), the fourth is also a function of Φ ; hence, the stronger the asymmetric effect, the higher the unconditional fourth moment. Such a feature helps capture a widely recognized characteristic of financial returns, i.e. high excess kurtosis. The coefficient of kurtosis, k , can be derived directly from equations (7) and (8) and turns out to be

$$k = \frac{3(1 - \alpha - \beta)^2(w^2 + \Phi^2) + (1 - \alpha - \beta)6w^2(\alpha + \beta)}{w^2(1 - \alpha^2 - 3\beta^2 - 2\alpha\beta)} \quad (9)$$

It is an increasing function of Φ , so that the Sign-switching ARCH model interprets high kurtosis also as the consequence of the asymmetric behaviour of the volatility (beyond its persistence).

³ As already recalled, equation (6) is related to the Sign-conditional AR(1) model reported in Granger and Teräsvirta (1993, pp. 137–9).

According to recent empirical evidence reported in Rabemananjara and Zakoian (1993), the Sign-switching ARCH model might be unsuccessful in detecting some of the non-linear characteristics of the volatility dynamics. In fact — as the authors point out — high negative shocks increase future volatility more than high positive ones while — at the same time — small positive shocks too often produce a stronger impact on future volatility than negative shocks of the same size. Thus, following the occurrence of a shock of a certain size, the asymmetric behaviour of the volatility might become reversed; the modeling of this feature is the focus of the remainder of the paper.

First, we define the size of the shock at which the reversal occurs, which also helps to clarify why the asymmetric behaviour of the volatility may eventually change direction. The ‘size measure’ which we employ in this paper is the level of unexpected volatility generated by a shock at time $t - 1$ (ε_{t-1}). Conditionally on the information set dated $t - 2$, the expected value of ε_{t-1}^2 is σ_{t-1}^2 . If, however, $\varepsilon_{t-1}^2 \geq \sigma_{t-1}^2$ ($\leq \sigma_{t-1}^2$), we shall say that ε_{t-1} has generated (at time $t - 1$) a level of volatility higher (lower) than expected (at time $t - 2$). Consider now a very small negative shock at time $t - 1$. If it produces a level of volatility at time $t - 1$ lower than expected at time $t - 2$, there should be no reason to believe that volatility at time t will increase as a consequence of the leverage effect. Roughly speaking, a small negative shock which generates lower volatility than expected may be regarded as good news; at the same time, positive shocks which generate lower volatility than expected may be regarded as relatively good news. This explains the mechanism according to which the reversals originate.

Let us now analyse more formally what may originate reversals. Black (1976) and Nelson (1991) observed that a negative shock on the stock of a given firm raises both its leverage ratio and its riskiness. As a result of this, the volatility of the stock — a measure of the riskiness of the firm — will increase as well; the opposite happens in the case of a positive shock. The two components — taken together — give rise to the leverage effect. However, it is worth noting that changes of the leverage ratio are likely to be followed by changes in the expected performance of the firm, the latter being a function of the differential between the expected average performance of the sector in which the firm operates and the overall cost of debt. Suppose that the economy has k productive sectors; thus

$$i_j = \rho_k + \pi_k \theta_j \quad (10a)$$

$$\pi_k \equiv \rho_k - r \quad (10b)$$

$$\theta_j \equiv D/S_j \quad (10c)$$

where i_j , S_j and θ_j are, respectively, the expected profitability, the price of the stock and the leverage ratio of the j th firm in the k th sector, r the interest paid on debt, ρ_k the average performance of the k th sector ($k = 1, \dots, K$) and D the amount of debt. Such a relation can be found in Modigliani and Miller (1958, proposition II, p. 271).

Consider the case that $\pi_k > 0$ in equation (10a). Then, a negative shock on S_j may be regarded as more favourable than a positive shock; in fact it increases θ_j and the expected profitability of the firm. However, if the negative shock is very large, two things may be hypothesized to happen: first, economic agents may discount a recession of the k th sector, i.e. a fall of ρ_k ; second the cost of the debt may be thought to start rising sharply for the j th firm, which happens when r is positively related to θ_j (hence inversely related to S_j).

Both events are likely to affect the sign of π_k hence causing — according to the explanation given above — reversals of the asymmetric reaction of the conditional volatility to the sign of past shocks. However, when the sign of π_k changes (and as long as it stays positive) negative

shocks of a modest size will induce less volatility than positive shocks of a similar size, according to equation (10a).

Past research has generally overlooked the impact of previous (unexpected or expected) volatility on its current expected level. Engle and Ng (1993), for example, propose to analyse the impact of news on the current conditional variance (i.e. on σ_t^2), keeping constant the information dated $t - 2$ and earlier, with all the lagged conditional variances evaluated at their unconditional value.

To define such issues formally, let $v_{t-1} = v_{t-1}(\varepsilon_{t-1})$ denote the (measurable) amount of unexpected volatility at time $t - 1$, generated by a shock occurred at time $t - 1(\varepsilon_{t-1})$. Let $f(v_{t-1})$ be some deterministic and measurable function mapping v_{t-1} onto the current conditional volatility. Then, if $g[\varepsilon_{t-1}, |\varepsilon_{t-1}|, \text{sign}(\varepsilon_{t-1})]$ is a deterministic, asymmetric and measurable response function of the current conditional volatility with respect to both size and sign of ε_{t-1} , other things equal, all the asymmetric ARCH models so far proposed in the literature focus mainly on modeling $g(\cdot)$ rather than $f(\cdot)$.

In order to take into account the impact of past unexpected volatility on future expected volatility, one has to build plausible functional forms for $f(v_{t-1})$, thus providing a model for the *response function of the future expected volatility to past unexpected volatility*; the latter would parallel the notion of ‘news impact curve’ of Engle and Ng (1993).

In this paper we will assume that $f(v_{t-1})$ is proportional to v_{t-1} . Consider, for example, the following model:

$$\sigma_t^2 = w + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + s_{t-1} v_{t-1} \quad (11a)$$

where s_{t-1} is defined by relations (5a–c) and

$$v_t \equiv \delta_0 \varepsilon_t^2 - \delta_1 \sigma_t^2 - \delta_2 \quad (11b)$$

so that v_t is a linear combination of the difference between the observed conditional volatility (ε_t^2) and its estimate, based on the available set of information, thus playing a role similar to that of an error-correcting variable. In the remainder of the paper, we will refer to model (11a,b) as Volatility-switching ARCH (henceforth VS).⁴

If $v_{t-1} < 0$, then, *ceteris paribus*, negative shocks generate more volatility than positive ones. However, if $v_{t-1} > 0$, positive shocks increase volatility more than negative ones. Thus, model

⁴ It is worth noting that the VS model (11a,b) generalizes the GJR, the latter being formulated as

$$\sigma_t^2 = w + \alpha \varepsilon_{t-1}^2 + \rho \sigma_{t-1}^2 + g_1 S_{t-1}^- \varepsilon_{t-1}^2 \quad (i)$$

where S_{t-1}^- is a dummy variable which takes the value +1 if ε_{t-1} is negative, and zero otherwise. Though in the VS case S_t is a dummy which takes on the values minus or plus one—instead of zero and one as in the GJR— S_{t-1}^- can always be written as

$$S_{t-1}^- = (\varepsilon_{t-1} - |\varepsilon_{t-1}|)(2\varepsilon_{t-1})^{-1} \quad (ii)$$

so that substituting equation (ii) into equation (i) and rearranging, one gets:

$$\sigma_t^2 = w + g_0 \varepsilon_{t-1}^2 + \rho \sigma_{t-1}^2 + \Phi s_{t-1} \varepsilon_{t-1}^2 \quad (iii)$$

where

$$g_0 \equiv \alpha + g_1/2 \quad (iv)$$

$$\Phi \equiv -(g_1/2) \quad (v)$$

Thus, the GJR model (iii)–(v) is obtained from the VS model (11a,b) when $\delta_0 = \Phi$ and $\delta_1 = \delta_2 = 0$.

(11a,b) is able to detect situations where the asymmetric behaviour of the volatility is reversed and further illustrates what has to be meant by *size* of a shock. *Small* shocks are those which produce a level of volatility lower than expected; *high* shocks are those which generate a level of volatility higher than expected.

The second and fourth moments of the innovations of the Volatility-switching model are (see Appendix 1):

$$E(\varepsilon^2) = w(1 - \alpha - \beta)^{-1} \quad (12a)$$

$$E(\varepsilon^4) = \frac{[(3w^2 - 3\delta_2^2)(1 - \alpha - \beta) + (6\alpha w^2 + 6\beta w^2 + 6\delta_1\delta_2 w - 6\delta_0\delta_2 w)]}{(1 - \alpha - \beta)(1 - 3\alpha^2 - \beta^2 - 2\alpha\beta - 3\delta_0^2 - \delta_1^2 - 2\delta_0\delta_1)} \quad (12b)$$

Since δ_0 , δ_1 and δ_2 measure the impact of v_{t-1} on σ_t^2 , deeper asymmetries in volatility will result in more leptokurtic distributions for the unconditional innovations.

2.2. Continuous Time Behaviour of the Models

The derivation of continuous time limits for the asymmetric models developed so far is obtained by increasingly partitioning time in equations (3)–(6) finely, according to the following scheme, where h denotes sampling frequency:

$${}^h\varepsilon_{hk} \equiv {}^h z_{hk} \cdot {}^h\sigma_{hk} \quad (13a)$$

$${}^h z_{hk} \sim N(0, h) \quad (13b)$$

$$s_k \equiv z_{hk} / |z_{hk}| \quad (13c)$$

$$s_k \sim \text{i.i.d.}(0, 1) \quad (13d)$$

$${}^h\sigma_{h(k+1)}^2 - {}^h\sigma_{hk}^2 = w_h + \delta_{2,h}s_k + {}^h\sigma_{hk}^2 [h^{-1}(\alpha_h - \delta_{0,h}s_k)z_{hk}^2 + (\beta_h - \delta_{1,h}s_k - 1)] \quad (13e)$$

Such a system is Markov, and we are interested in analysing the conditions under which it converges weakly (i.e. in distribution) to an Itô diffusion process, as h drops to zero. To do this, we retain Nelson's (1990) assumptions (1 to 5), omitting, for simplicity, more general issues, such as the conditions under which stochastic difference equations converge to stochastic differential equations.

Before deriving the diffusion limits of the models proposed so far, we report also analogous results for the Power ARCH model of Ding *et al.* (1993), since it has quite a general structure and encompasses many heteroscedastic formulations, including the GJR. The latter — as reported by Engle and Ng (1993) — seems the best parametric model to capture the asymmetry of volatility and therefore its empirical performance will be compared to that of the Volatility-switching model.

To start with, let us replace equations (6) and (13e) with the respective Power ARCH equations:

$$\sigma_{t+1}^\delta = w' + \alpha'(|\varepsilon_t| - \tau\varepsilon_t)^\delta + \beta\sigma_t^\delta \quad (6')$$

$$({}^h\sigma_{h(k+1)}^{(h)})^\delta - ({}^h\sigma_{hk}^{(h)})^\delta = w'_h + [\beta_h + h^{-0.5\delta} |{}^h z_{hk}^{(h)}|^\delta (1 - \tau s_k)^\delta \alpha'_h - 1] ({}^h\sigma_{hk}^{(h)})^\delta \quad (14)$$

Fornari and Mele (1995a) showed that the diffusion limit of equation (14) is:

$$d\sigma_t^\delta = (w - \theta_\delta \sigma_t^\delta) dt + \Omega_\delta \sigma_t^\delta dW_t \quad \theta \geq 0 \quad (15)$$

where W_t is a standard (scalar) Brownian motion, w is the continuous time counterpart (*c.t.c.*) of w' , and

$$\theta_\delta = c.t.c. \alpha' \frac{2^{0.5(\delta-1)} \Gamma[0.5(\delta+1); 0.5]}{(2\pi)^{0.5}} [(1+\tau)^\delta + (1-\tau)^\delta] + \beta - 1 \quad (16)$$

$$\begin{aligned} \Omega_\delta^2 = c.t.c. \alpha' & \frac{2^{(\delta-0.5)} \Gamma[(\delta+0.5); 0.5]}{(2\pi)^{0.5}} [(1+\tau)^{2\delta} + (1-\tau)^{2\delta}] \\ & - \frac{2^{0.5(\delta-1)} \Gamma[0.5(\delta+1); 0.5]}{(2\pi)^{0.5}} [(1+\tau)^\delta + (1-\tau)^\delta]^2 \end{aligned} \quad (17)$$

where $\Gamma[0.5(\delta+1); 0.5] \equiv \int_{0,\infty} (0.5)^{0.5(\delta+1)} X^{0.5(\delta-1)} \exp(-0.5X) dX$.

It is easy to check that the diffusion limit (15) collapses to Nelson's (1990) standard result for the GARCH(1,1) when $\delta = 2$ and $\tau = 0$; in the case of the GJR model, which is nested into the Power ARCH, equation (6') reduces to

$$\sigma_t^2 = w + g_0 \varepsilon_{t-1}^2 + p \sigma_{t-1}^2 + \Phi s_{t-1} \varepsilon_{t-1}^2 \quad (18)$$

when (Ding *et al.*, 1993)

$$\delta = 2 \quad (19a)$$

$$\alpha = \alpha'(1-\tau)^2 \quad (19b)$$

$$g_1 = 4\alpha'\tau \quad (19c)$$

Substitution of relations (19a-c) into (16) and (17) gives, after tedious but straightforward algebra:

$$\theta_2^{\text{GJR}} = c.t.c. p + \alpha + (g_1/2) - 1 \quad (20a)$$

$$\Omega_2^{\text{GJR}} = c.t.c. 2g_0^2 + 0.75g_1^2 \quad (20b)$$

suggesting that the diffusion limit of the GJR model has the following form:

$$d\sigma_t^2 = (w - \theta_2^{\text{GJR}} \sigma_t^2) dt + \Omega_2^{\text{GJR}} \sigma_t^2 dW_t \quad (21)$$

where W_t is a standard (scalar) Brownian motion.⁵

At this point it is important to investigate whether the diffusion limit (15) is able to generalize the diffusion of the Sign- and Volatility-switching ARCH models; in this case, in fact, they would be merely particular cases of the Power ARCH. To anticipate the results, it turns out that the answer is negative for both models.

To show this, let us start with the VS model. We first evaluate the expected value per unit of time of $({}_h\sigma_{(k+1)}^2 - {}_h\sigma_{hk}^2)$ in equation (13e), with σ_h^2 generated by equation (11a). To avoid the explosion of the drift per unit of time as $h \rightarrow 0$, we require the following Lipschitz conditions:

$$\lim_{h \rightarrow 0} h^{-1} w_h = w \quad (22)$$

$$\lim_{h \rightarrow 0} h^{-1} (\alpha_h + \beta_h - 1) \equiv \lim_{h \rightarrow 0} h^{-1} \theta_h = -\theta \quad \theta \geq 0 \quad (23)$$

⁵ In Glosten *et al.* (1993) the persistence of the conditional variance is computed by regressing the (estimated) σ_t^2 on its own lag, σ_{t-1}^2 . This procedure, however, lacks a rigorous theoretical justification. Relation (20b), instead, provides the analytical expression for the amount of persistence of the GJR model. Note that this is an increasing function of g_1 .

obtaining

$$\begin{aligned} \lim_{h \rightarrow 0} \{E[h^{-1}({}_h\sigma_{h(k+1)}^2 - {}_h\sigma_{hk}^2) | F_{hk}]\} &= \lim_{h \rightarrow 0} [h^{-1}w_h + h^{-1}(\alpha_h + \beta_h - 1){}_h\sigma_{hk}^2] \\ &= w - \theta\sigma_t^2 \end{aligned} \quad (24)$$

The evaluation of $E[h^{-1}({}_h\sigma_{h(k+1)}^2 - {}_h\sigma_{hk}^2) | F_{hk}]^2$ gives:

$$\begin{aligned} E[h^{-1}({}_h\sigma_{h(k+1)}^2 - {}_h\sigma_{hk}^2)^2 | F_{hk}] &= h^{-1}E\{w_h^2 + \delta_{2,h}^2 s_k^2 + (\alpha_h + \beta_h - 1)^2 {}_h\sigma_{hk}^4 \\ &+ [2(\alpha_h^2 + \delta_{0,h}^2) + (\delta_{0,h} - \delta_{1,h})^2] {}_h\sigma_{hk}^4 + 2w_h(\beta_h + \alpha_{hh} - 1){}_h\sigma_{hk}^2\} \end{aligned} \quad (25)$$

Using the Lipschitz conditions (22) and (23) and assuming the existence of the following limits:

$$\lim_{h \rightarrow 0} h^{-1} \delta_{2,h}^2 = \delta_2^2 \quad (26)$$

$$\lim_{h \rightarrow 0} h^{-1} [2(\alpha_h^2 + \delta_{0,h}^2) + (\delta_{0,h} - \delta_{1,h})^2] = \Phi^2 \quad (27)$$

we get:

$$\lim_{h \rightarrow 0} E[h^{-1}({}_h\sigma_{h(k+1)}^2 - {}_h\sigma_{hk}^2) | F_{hk}]^2 = \delta_2^2 + \Phi^2 \sigma_t^4 \quad (28)$$

Relations (24)–(28) suggest that $\{{}_h\sigma_{h(k+1)}^2\}_{k=0,\infty}$ converges in distribution to a diffusion limit of the following form, as $h \rightarrow 0$:

$$d\sigma_t^2 = (w - \theta\sigma_t^2) dt + (\delta_2^2 + \Phi^2 \sigma_t^4)^{0.5} dW_t \quad \theta \geq 0 \quad (29)$$

where dW_t denotes the increments of a (scalar) standard Brownian motion. It is straightforward to check that the structural form of such diffusion limit equals the same expression as in Nelson (1990, eq. 2.40), when δ_0 , δ_1 and $\delta_2 = 0$ and to derive the diffusion limit of the Sign-switching model (3)–(6), which is essentially the same as equation (29) once δ_0 and δ_1 have been set equal to nil in equations (25) and (27).

2.3. The Stationary Distribution of the Conditional Variance

With the GARCH(1,1), Nelson (1990) showed that the stationary distribution of the conditional variance becomes an inverted Gamma as h (the sampling frequency) drops to zero. Then it is interesting to analyse what this distribution is in the case of both the Sign- and Volatility-switching models. In Appendix 2, we show that such a distribution is:

$$P(v | v_0) \text{ proportional to: } (r^2 + v^2)^{-m} e^{-n \cdot \text{arctog}(qv)} \quad (30)$$

where $v \equiv \sigma^2$, v_0 the initial condition, $r^2 \equiv (\theta/\alpha)^2$, $m \equiv (\theta/\alpha^2) + 1$, $n = 2w|\Phi|^{-1}$, $q \equiv \alpha(|\Phi|)^{-1}$. Note that

$$\lim_{\Phi \rightarrow 0+} P(v | v_0) = \lim_{\Phi \rightarrow 0+} -P(v | v_0) \text{ proportional to } v^{-2m} e^{-k/v} \quad (31)$$

where $k \equiv 2w\alpha^{-2}$. Distribution (31) is an inverted Gamma, consistent with Nelson's (1990) results.

From equation (30) it is easy to find the stationary distribution of the standard deviation of the innovations of the Sign- and Volatility-switching models, which is

$$f(\sigma | \sigma_0) d\sigma \text{ proportional to: } 2\sigma(r^2 + \sigma^4)^{-m} e^{-n \cdot \text{arctog}(q\sigma^2)} d\sigma \quad (32)$$

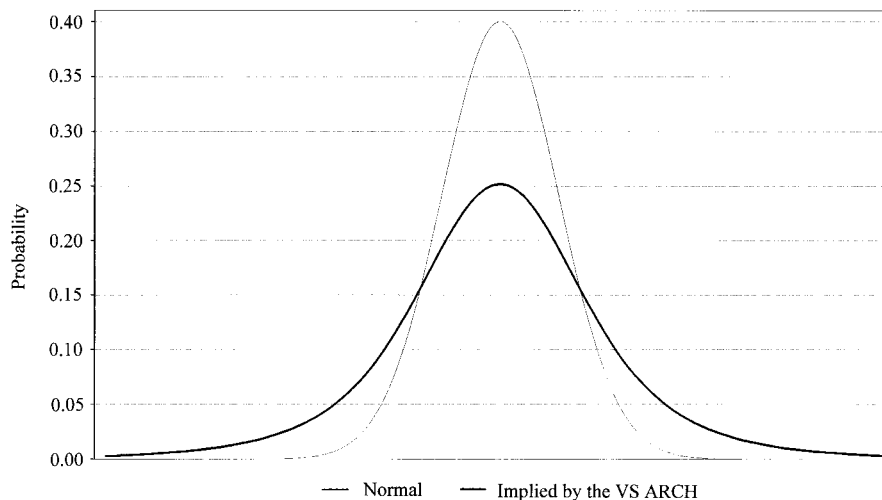


Figure 1. Stationary distribution of residuals

Relation (32) implies that the innovations of the Sign-switching or, alternatively, Volatility-switching model, say ε^* , have a stationary density function, $p(\varepsilon^*)$, which solves:

$$p(\varepsilon^*) = \int_{0,\infty} N(\varepsilon^* \sigma^{-1}) f(\sigma) \sigma^{-1} d\sigma \quad (33)$$

where $N(\cdot)$ is a standard normal density function. Unfortunately we did not manage to solve the integral in equation (33) analytically; hence, a numerical procedure—the adaptive recursive Simpson’s rule—was employed (see e.g. Abramowitz and Stegun, 1970, formula (25.4.5), p. 866). Results are reported in Figure 1, where the density $p(\varepsilon^*)$ defined in equation (33) is compared to a standard normal density.⁶ It is easy to note that the former has fatter tails compared to the corresponding area of the normal variate; such a feature helps to capture the high number of outliers observed in empirical distributions.

3. EMPIRICAL ANALYSIS

We have employed six stock market indices to test and compare the empirical performances of the Sign- and Volatility-switching ARCH models to the GJR formulation; the latter is chosen as a benchmark since—as reported by Engle and Ng (1993)—it represents the best parametric model for asymmetric conditional variances. The series are the Standard and Poor’s 500 (United States), Topix (Japan), CAC40 (France), FT-100 (United Kingdom), FAZ (Germany) and MIB (Italy) indices, observed daily from 1 January 1990 to 16 October 1995. The sample includes 1494 observations.

Let P_t^j denote the level of the j th stock index at time t . Then we evaluate six series of *unpredictable* returns, u_t , which are the residuals of univariate regressions of *ex-post* returns

⁶ All the computations were performed with the Quadrature Routine of Matlab. The values of the parameters used in the numerical integration procedure were $w = 0.12$, $\alpha = 0.15$, $\theta = 0.0045$, $\Phi = 0.010$. The range of variation for ε^* was $[-8, +8]$.

r_t^j [$r_t^j = \log(P_t^j/P_{t-1}^j)$] on a constant and on their lagged values, i.e.

$$r_t^j = \mu_0^j + \sum_{i=1, \dots, p} \mu_i^j r_{t-i}^j + u_t^j \quad (34)$$

where μ_0^j and μ_i^j ($i = 1, \dots, p$) are real, non-stochastic parameters, i is a suitably chosen lag and $j = 1, \dots, 6$ denotes the six markets. In all the regressions i was chosen to be one according to the criterion of Schwarz (1978).

Table I. Preliminary tests on the unpredictable returns^a

| Country→ Test↓ | United States | Germany | Japan | United Kingdom | France | Italy | Critical values |
|--------------------------|------------------|---------|---------|-------------------|--------|--------|--------------------|
| Q_5 | 3.03 | 6.94 | 9.62 | 12.61 | 10.66 | 27.56 | 11.10 |
| Q_{10} | 18.00 | 19.04 | 18.51 | 27.19 | 11.87 | 37.41 | 18.30 |
| TR_5^2 | 73.00 | 111.52 | 142.75 | 39.52 | 67.81 | 55.29 | 11.10 |
| SKEW | -0.05 | 0.10 | 0.34 | 0.38 | -0.06 | -0.14 | 0 |
| KU | 3.53 | 3.21 | 5.84 | 4.14 | 3.02 | 2.42 | 0 |
| JB | 773.12 | 640.16 | 2151.17 | 1101.99 | 568.56 | 367.69 | 5.99 |
| SBT | 0.39 | 0.57 | 1.31 | -0.28 | -0.28 | -0.24 | > 2 or < -2 |
| NSBT | -2.03 | -3.59 | -6.78 | 0.21 | -2.47 | -3.55 | > 2 or < -2 |
| PSBT | 4.99 | 0.99 | 2.65 | 4.28 | 1.31 | 0.12 | > 2 or < -2 |
| Joint (TR ²) | 46.81 | 19.46 | 69.08 | 26.38 | 13.57 | 21.98 | 7.81 |

^a Q_5 and Q_{10} are the Box and Pierce Q -tests up to 5 and 10 lags, respectively; TR_5^2 the Engle's TR^2 computed up to the fifth lag; SKEW is the coefficient of skewness; KU the coefficient of excess kurtosis; JB is the Jarque and Bera test, SBT is the Sign Bias test; NSBT is the Negative Sign Bias test; PSBT is the positive Sign Bias test; JOINT (TR²) is the Lagrange Multiplier test for the null that the squared unpredictable returns are not explained by the lagged values of S^- , S^-u and S^+u .

Table I shows a number of preliminary statistics for the six unpredictable returns. The presence of autocorrelation for the residuals of equation (34) was ascertained by means of the Box and Pierce's Q -test evaluated up to the fifth and tenth lags. Under the null of no autocorrelation, such statistics are asymptotically distributed as chi-squares with five and ten degrees of freedom, respectively; their 5% critical levels are 11.1 and 18.3. The hypothesis of autocorrelation in the second-order moments has been tested via the TR^2 (Engle, 1982); it is evaluated as the TR^2 of the regression of the squared residuals of equation (34) on a constant and five own lags and is asymptotically distributed as a chi-square with five degrees of freedom. As far as the shape of the unconditional distribution of the unpredictable returns is concerned, we report the coefficients of skewness and kurtosis, along with the Jarque and Bera (JB) test. The latter is asymptotically distributed as a chi-square with two degrees of freedom, under the null that the data come from a normal distribution. Its 5% critical level is 5.99.

The tests for the presence of asymmetric behaviour of the volatility developed by Engle and Ng (1993) are also performed. These are the Sign Bias test (SBT), the Negative Sign Bias test (NSBT), the Positive Sign Bias Test (PSBT), and the Joint test (TR²). SBT, NSBT and PSBT are the t -statistics for the coefficients of a linear regression of the squared innovations of regression (34) on S_{t-1}^- , $S_{t-1}^-u_{t-1}$, and on $S_{t-1}^+u_{t-1}$, respectively, with S_{t-1}^- being a dummy variable which equals plus one if $\text{sign}(u_t) = -1$, and zero otherwise and $S_{t-1}^+ = 1 - S_{t-1}^-$. The three tests can also be run jointly as the TR^2 of the following regression:

$$u_t^2 = c_0 + c_1 S_{t-1}^- + c_2 S_{t-1}^- u_{t-1} + c_3 S_{t-1}^+ u_{t-1} + z_t \quad (35)$$

where c_i ($i = 0, 1, 2, 3$) are real, non-stochastic parameters and z_t is a white-noise process. Such a statistic is asymptotically distributed as a chi-square with three degrees of freedom; its critical 5% threshold is 7.81.

All the unpredictable returns are not significantly autocorrelated, with the exception of the UK and Italy. Their unconditional distributions are not normal since the coefficients of skewness and the kurtoses diverge from the values typical of a Gaussian distribution — zero and three — and the JB test rejects the normality hypothesis at any reasonable level of confidence. The evidence for ARCH effects is clearly supported by the TR_3^2 . Asymmetries in volatility exist in all the series, according to the joint test. Coming to the estimation, the GJR(1,1) models have the following structure:

$$r_t = \mu_0 + \mu_1 r_{t-1} + u_t \quad (36a)$$

$$u_t | I_{t-1} \sim N(0, \sigma_t^2) \quad (36b)$$

$$\sigma_t^2 = w + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 + \Phi_1 s_{t-1} u_{t-1}^2 \quad (36c)$$

where $\mu_0, \mu_1, w, \alpha, \beta$ and Φ_1 are real, non-stochastic parameters, r_t is the *ex-post* return of the j th index ($j = 1, \dots, 6$), s_t is plus one if the sign of the forecast error dated t — i.e. $\text{sign}(u_t)$ — is negative and zero otherwise. The structural form of the Volatility Switching Model is the same as equations (36a–c) but (36c) is replaced with:

$$\sigma_t^2 = w + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 + s_{t-1} v_{t-1} \quad (36d)$$

where:

$$v_t \equiv \delta_0 u_t^2 - \delta_1 \sigma_t^2 - \delta_2 \quad (36e)$$

and s_t is plus one if the sign of u_t is positive, minus one if it is negative. Table II shows the values of the parameters of the GJR models along with their t -ratios and Table III those of the VS; the Sign-switching models have not been estimated since — if they outperformed the VS — the null hypothesis that $\delta_0 = \delta_1 = 0$ would not be rejected, which never happens in the analysed cases.

Table II. Parameters of the GJR models (t -ratios in parentheses)

| Country→ Parameter↓ | United States | Germany | Japan | United Kingdom | France | Italy |
|------------------------|---------------------|---------------------|-----------------------|---------------------|-----------------------|-----------------------|
| μ_0 | 2.42E – 7 (1.28) | 3.17E – 5 (0.14) | –6.07E – 4 (–2.38) | 1.23E – 4 (0.68) | –2.25E – 4 (–0.79) | –1.94E – 4 (–0.63) |
| μ_1 | 0.34 (1.14) | 0.01 (0.24) | 0.14 (5.06) | 0.07 (2.76) | 2.84E – 2 (0.99) | 0.25 (8.68) |
| w | 5.26E – 6 (8.26) | 6.92E – 6 (9.25) | 8.82E – 6 (8.37) | 2.68E – 6 (4.82) | 1.74E – 5 (6.46) | 7.35E – 6 (6.43) |
| α | 0.09 (18.28) | 0.11 (16.37) | 0.16 (21.16) | 0.09 (13.33) | 0.07 (9.06) | 0.06 (14.51) |
| β | 0.81 (94.94) | 0.82 (95.04) | 0.80 (94.94) | 0.87 (86.32) | 0.80 (49.84) | 0.89 (131.42) |
| ϕ_1 | –0.04 (4.57) | –0.07 (5.79) | –0.10 (–8.48) | –0.04 (–3.85) | –0.06 (–5.29) | –0.04 (–4.06) |
| Log of likelihood | 6634.78 | 6284.86 | 5938.07 | 6640.39 | 5967.58 | 5925.13 |

Table III. Parameters of the volatility-switching models (*t*-ratios in parentheses)

| Country→ Parameter↓ | United States | Germany | Japan | United Kingdom | France | Italy |
|---------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| μ_0 | 1.71E - 4 (0.95) | 3.88E - 6 (0.02) | -5.71E - 4 (-2.15) | 8.57E - 5 (0.46) | -1.88E - 4 (-0.69) | -2.85E - 4 (-0.98) |
| μ_1 | 0.05 (1.72) | -0.01 (-0.59) | 0.15 (5.12) | 0.08 (2.82) | 0.03 (1.08) | 0.26 (8.82) |
| w | 5.29E - 6 (8.48) | 9.06E - 6 (8.13) | 6.60E - 6 (6.54) | 2.78E - 6 (4.81) | 9.60E - 6 (9.31) | 1.02E - 5 (6.63) |
| α | 0.09 (13.83) | 0.12 (14.48) | 0.15 (18.09) | 0.09 (13.08) | 0.05 (10.08) | 0.11 (12.99) |
| β | 0.81 (79.54) | 0.79 (67.09) | 0.83 (98.88) | 0.87 (86.54) | 0.87 (125.73) | 0.83 (75.39) |
| δ_0 | -0.06 (-5.64) | -0.07 (-4.40) | -0.09 (-5.90) | -0.03 (-2.69) | -0.05 (-4.31) | -0.08 (-5.14) |
| δ_1 | -0.11 (-2.58) | -0.01 (-0.22) | 0.20 (5.80) | 0.06 (2.32) | 0.19 (2.47) | 0.15 (3.07) |
| δ_2 | 6.57E - 6 (3.63) | 2.68E - 6 (0.66) | -1.46E - 5 (-4.80) | -1.39E - 6 (-0.74) | -1.31E - 5 (-1.56) | -2.47E - 5 (-5.42) |
| Log of likelihood | 6679.75 | 6285.57 | 5946.78 | 6642.53 | 5973.41 | 5928.78 |
| χ^2_2 of VS against GJR | 89.94 (prob = 0.0) | 1.42 (prob = 49.2) | 8.71 (prob = 1.3) | 4.28 (prob = 3.8) | 11.66 (prob = 0.3) | 7.30 (prob = 2.6) |

Table IV. Tests on residuals standardized with the standard deviation of the volatility-switching model

| Country→ Test↓ | United States | Germany | Japan | United Kingdom | France | Italy |
|-------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Kurtosis ^a | 2.18 2.43 2.65 | 1.24 2.31 3.20 | 2.76 2.89 5.82 | 2.39 2.38 4.18 | 0.97 1.07 3.03 | 0.96 1.11 2.38 |
| Sign Bias test | 1.03 | 1.18 | -0.98 | 1.38 | -0.12 | 1.19 |
| Negative Sign Bias test | -0.04 | 0.06 | 0.52 | -0.16 | 0.46 | -1.46 |
| Positive Sign Bias test | -0.57 | 0.46 | 1.45 | -0.02 | -0.15 | -0.03 |

^a The second and third lines show, for comparison, the excess kurtosis of the residuals standardized with the GJR conditional standard deviation and the sample standard deviation, respectively.

Our empirical results confirm the outcome of the preliminary tests. All the series display both high ARCH and asymmetric effects⁷ as evidenced by the significance of α , β , ϕ_1 , δ_0 and δ_1 . However, based on a likelihood ratio test reported in the last line of Table III the Volatility-switching model fits the patterns of the data better than the GJR in all cases except Germany, where reversals of the asymmetric reaction of the conditional volatility to the sign of past news do not exist, at least over the sample taken into consideration. The Volatility-switching model outperforms the GJR also when the models are ranked according to the kurtosis of the residuals standardized with the respective conditional standard deviations, except for the UK, where the two kurtoses coincide. In both cases, however, the models capture successfully the asymmetric behaviour of the conditional volatility. This is highlighted by the tests of Engle and Ng (1993)—reported in Table IV only for the Volatility-switching—performed on the squared residuals of

⁷ In the estimated models, the lags chosen for the conditional mean equation (36a) coincide with those employed in the preliminary analysis. The likelihood has been maximized by means of the Berndt *et al.* (1974) procedure.

equation (36a) standardized with the conditional standard deviation. In all cases the SBT, PSBT and NSBT tests do not reveal any presence of asymmetry in the reaction of such series to the sign of past forecast errors. In all cases, the generating processes of the conditional variances are stationary; in fact their persistence — as roughly measured by $\alpha + \beta$ — is always below unity.

Despite a similar empirical performance, what distinguishes the two models is the interpretation of the reversal of the asymmetric behaviour of the conditional volatility. Its existence is highlighted by the significance of δ_0 and δ_1 for all countries, except Germany, in the VS model while it is ignored by the GJR, which gives it *a priori* a probability equal to nil. To clarify this point denote, assuming that $\delta_0 < 0$ which is the case of our estimates,

$$\Pr\{\text{reversal}\} = \Pr\{\varepsilon_{t-1}^2 < k_0 E(\varepsilon_{t-1}^2 | I_{t-2}) + k_1\} = \Pr\{\varepsilon_{t-1}^2 < k_0 \sigma_{t-1}^2 + k_1\}$$

where $k_0 = \delta_1/\delta_0$, $k_1 = \delta_2/\delta_0$, I_{t-2} is the information set available to economic agents when the expectations are formed and $E(\cdot)$ is the conditional expectation operator. Dividing the above relation by σ_{t-1}^2 and recalling that $\varepsilon_t^2/\sigma_t^2$ is distributed as a chi-square with one degree of freedom, we obtain

$$\Pr\{\text{reversal}\} = \Pr\{\chi_1^2 < k_0 + k_1 \sigma_t^{-2}\}$$

Note that the probability of reversal is an increasing function of the conditional precision process, when $\delta_2 < 0$. Then, if $(\delta_1 \geq 0 \cap \delta_2 \geq 0)$ holds, the probability of reversal is nil; such a situation arises for the GJR which — as recalled in footnote 4 — is a special case of a Volatility-switching once $\delta_1 = \delta_2 = 0$. Thus the Volatility-switching model allows a more detailed analysis of the asymmetric behaviour of the volatility enabling, by construction, shifts in its direction, according to the size of past shocks.

To conclude, it is worth noting that the estimated models have a disturbing feature, the remedy for which is outside the aim of this paper. What happens is that although the kurtosis of the residuals standardized with the conditional standard deviation is lower than in the original series, it is still higher than 3, the value typical of a normal variate. At a very least, this suggests that the distributional assumptions made for the conditional distribution of the innovations should be modified. It may turn out that other hypotheses, such as the Generalized Error distribution, or the Student- t , could improve the results achieved so far.

4. CONCLUSIONS

Sign- and Volatility-switching models have been presented in the paper. They allow for an asymmetric behaviour of the conditional volatility with respect to negative and positive shocks, since they map the sign of past forecast errors onto the current, conditional volatility. The Volatility-switching model is also able to capture *reversions* in the asymmetric behaviour of the volatility. Weak convergence results presented for both models show that they converge in distribution to Itô diffusion processes; also they are shown to have a stationary distribution function, for which a closed form solution was provided. Numerical procedures were employed to compute the stationary distribution of the innovations of these models. Empirical analysis has shown that the proposed models provide better interpretative results than the GJR, since they are able to give further insights regarding the reversion of the asymmetry of the conditional volatility.

Two issues deserve further research: first, the hypothesis of conditional normality for the innovations of the model needs to be modified; second, the response function of the

Volatility-switching model is just a linear function of v_{t-1} , the size of previous unexpected volatility, times the sign of past forecast errors. Other (non-linear) specifications might be plausible, although the weak convergence analysis for such specifications would inevitably become more intricate.

APPENDIX 1

Unconditional Moments of ε for the Sign-switching GARCH(1,1,1)

Given that $\varepsilon_t | I_{t-1} \sim N(0, \sigma_t^2)$, and:

$$E(\varepsilon_t^{2m} | I_{t-1}) = (w + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2 + \Phi s_{t-1})^m h_{2m} \quad (\text{A1})$$

where $h_{2m} \equiv \prod_{i=1, \dots, m} (2j - 1)$. Setting $m = 1$ in equation (A1) yields:

$$E(\varepsilon_t^2 | I_{t-1}) = w + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2 + \Phi s_{t-1}$$

when t is allowed to tend to infinity, so that the dependence of the current values on their past realization becomes negligible, we get:

$$E(\varepsilon^2) = w + \beta E(\sigma^2) + \alpha E(\varepsilon^2) + \Phi E(s) = w/(1 - \beta - \alpha) \quad (\text{A2})$$

Setting $m = 2$ and letting time go to infinity in equation (A1) gives:

$$\begin{aligned} E(\varepsilon^4) &= 3E[w^2 + \beta^2\sigma^4 + \alpha^2\varepsilon^4 + \Phi^2 + 2\alpha\beta\sigma^2\varepsilon^2 + 2\beta w\sigma^2 + 2\alpha w\varepsilon^2] \\ &= 3w^2 + 3\beta^2 E(\sigma^4) + 3\alpha^2 E(\varepsilon^4) + 3\Phi^2 + 6\alpha\beta E(\sigma^4) + 6\beta w E(\sigma^2) + 6\alpha w E(\varepsilon^2) \\ &= 3w^2 + 3\beta^2 E(\sigma^4)/3 + 3\alpha^2 E(\varepsilon^4) + 3\Phi^2 + 6\alpha\beta(1/3)E(\sigma^4) + 6\beta w(w/(1 - \beta - \alpha)) \\ &\quad + 6\alpha w(w/(1 - \alpha - \beta)) \\ &= [3(w^2 + \Phi^2)(1 - \beta - \alpha) + 6w^2(\alpha + \beta)][(1 - \beta - \alpha)(1 - \beta^2 - 3\alpha^2 - 2\beta\alpha)]^{-1} \end{aligned} \quad (\text{A3})$$

The last expression collapses to the standard result of Bollerslev (1986) when $\Phi = 0$.

Unconditional Moments of ε for the Volatility-switching Model

Define the Volatility-switching model as

$$\sigma_t^2 = w + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + (\delta_0\varepsilon_{t-1}^2 - \delta_1\sigma_{t-1}^2 - \delta_2)s_{t-1}$$

and let ε_t be zero mean conditionally normally distributed, with conditional variance σ_t^2 . Hence:

$$E(\varepsilon_t^{2m} | I_{t-1}) = (w + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2 + s_{t-1}v_{t-1})^m h_{2m}$$

where $h_{2m} \equiv \prod_{j=1, \dots, m} (2j - 1)$. Setting $m = 1$ in equation (A2) gives:

$$E(\varepsilon_t^2 | I_{t-1}) = w + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + s_{t-1}v_{t-1}$$

Recursive substitution yields:

$$E(\varepsilon^2) = w/(1 - \beta - \alpha) \quad (\text{A4})$$

Setting $m = 2$, and letting time go to infinity in equation (A2), so that the dependence of current values on the past is negligible, gives:

$$\begin{aligned} E(\varepsilon^4) &= 3E[w^2 + \alpha^2\varepsilon^4 + \beta^2\sigma^4 + \delta_0^2\varepsilon^4 + \delta_1^2\sigma^4 + \delta_2^2 + 2\delta_0\delta_1\varepsilon^2\sigma^2 - 2\delta_0\delta_2\varepsilon^2 \\ &\quad + 2\delta_1\delta_2\sigma^2 + 2\alpha w\varepsilon^2 + 2\beta w\sigma^2 + 2\alpha\beta\varepsilon^2\sigma^2] \\ E(\varepsilon^4)[1 - 3\alpha^2 - \beta^2 - 3\delta_0^2 - \delta_1^2] &= (3w^2 - 3\delta_2^2) \\ &\quad + [-6\delta_0\delta_2w + 6\delta_1\delta_2w + 6\alpha w^2 + 6\beta w^2]/(1 - \alpha - \beta) \end{aligned}$$

so that

$$E(\varepsilon^4) = \frac{(3w^2 - 3\delta_2^2)(1 - \alpha - \beta) + [6\alpha w^2 + 6\beta w^2 - 6\delta_0\delta_2w + 6\delta_1\delta_2w]}{[1 - 3\alpha^2 - \beta^2 - 3\delta_0^2 - \delta_1^2](1 - \alpha - \beta)} \quad (\text{A5})$$

APPENDIX 2

The Stationary Distribution of σ^2 in the Sign- and Volatility-Switching Model

Let $v \equiv \sigma^2$, and rewrite equation (29) as:

$$dv = (w - \theta v) dt + (\Phi^2 + \alpha^2 v^2)^{0.5} dW \quad (\text{B1})$$

Let $p(s, t, v(t) | v_0)$ be the probability density function of v , given v_0 and $0 < s < t < \infty$. The Fokker–Planck–Kolmogorov forward diffusion equation (see e.g. Papoulis, 1965) associated with equation (B1) is:

$$\partial(p(s, t, v | v_0))/(\partial t) = 0.5\partial^2\{[(\Phi^2 + \alpha^2 v^2)p(s, t, v | v_0)]\}/(\partial v^2) - \partial\{[(w - \theta v)p(s, t | v_0)]\}/(\partial v)$$

where $\partial(\cdot)$ denotes the derivative operator.

As Nelson (1990) remarks (see also Papoulis, 1965; Wong, 1971), an invariant density function (given it exists) must satisfy:

$$0.5\partial\{v[(\Phi^2 + \alpha^2 v^2)p(v | v_0)]\}/(\partial v) = (w - \theta v)p(v | v_0) \quad (\text{B2})$$

where $p(v | v_0) = \lim_{t \rightarrow 0} p(s, t, v | v_0)$. Developing equation (B2) explicitly yields:

$$[(\partial P)/(\partial v)]/P = 2w(\Phi^2 + \alpha^2 v^2)^{-1} - 2(\theta + \alpha^2)v(\Phi^2 + \alpha^2 v^2)^{-1}$$

Hence

$$\ln(P) \propto -2w(\alpha|\Phi|)^{-1} \operatorname{arctg}(\alpha v/|\Phi|) - (\theta + \alpha^2)\alpha^{-2} \ln(\Phi^2\alpha^{-2} + v^2) \quad (\text{B3})$$

Taking the exponent of both sides of equation (B3) gives directly the density in equation (32) of the text.

ACKNOWLEDGEMENTS

We wish to thank Michel Botomazava, Giorgio De Santis, Andrew Harvey, Giuseppe Schlitzer and—above all—Mark Watson and two anonymous referees for valuable suggestions. The usual disclaimer applies.

REFERENCES

- Abramowitz, M. and I. Stegun (1970), *Handbook of Mathematical Functions*, Dover Publications, New York.
- Bera, A. and M. Higgins (1993), 'ARCH models: properties, estimation and testing', *Journal of Economic Surveys*, **7**, 305–66.
- Berndt, E., B. Hall, R. Hall and J. Hausman (1974), 'Estimation and inference in nonlinear structural models', *Annals of Economic and Social Measurement*, **4**, 653–65.
- Billingsley, P. (1968), *Convergence of Probability Measures*, Wiley, New York.
- Black, F. (1976), 'Studies of stock market volatility changes', in American Statistical Association (ed.), *Proceedings of the American Statistical Association: Business and Economic Statistics Section*, Chicago.
- Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics*, **31**, 307–27.
- Bollerslev, T., R. Chou and K. Kroner (1992), 'ARCH modeling in finance: a review of the theory and empirical evidence', *Journal of Econometrics*, **52**, 5–59.
- Campbell, J. and L. Hentschel (1992), 'No news is good news. An asymmetric model of changing volatility in stock returns', *Journal of Financial Economics*, **31**, 281–318.
- De Santis, G. (1991), 'Fitting the EGARCH model to the Italian stock market', *Ricerche Economiche*, **45**, 21–55.
- De Santis, G. and A. Sbordone (1990), 'A CAPM with a multivariate generalised ARCH process: an empirical analysis of the Italian financial market', University of Chicago, mimeo.
- Ding, Z., R. Engle and C. Granger (1993), 'A long memory property of stock market returns and a new model', *Journal of Empirical Finance*, **1**, 83–106.
- El Babsiri, M. and J. M. Zakoian (1994), 'Continuous time approximation of threshold GARCH models', CREST-ENSAEE, mimeo.
- Engle, R. F. (1982), 'Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation', *Econometrica*, **50**, 987–1008.
- Engle, R. F., D. Lilien and R. Robins (1987), 'Estimating time varying risk premia in the term structure: the ARCH-M model', *Econometrica*, **55**, 391–407.
- Engle, R. F. and C. Mustafa (1992), 'Implied ARCH models from option prices', *Journal of Econometrics*, **52**, 289–311.
- Engle, R. F. and V. Ng (1993), 'Measuring and testing the impact of news on volatility', *Journal of Finance*, **43**, 1749–77.
- Engle, R. F., V. Ng and M. Rothschild (1990), 'Asset pricing with a factor ARCH covariance structure: empirical estimates for Treasury bills', *Journal of Econometrics*, **45**, 213–38.
- Fornari, F. and A. Mele, (1995a), 'Weak convergence results for a general class of nonlinear ARCH models', *Econometric Reviews*, forthcoming.
- Fornari, F. and A. Mele (1995b), 'Continuous time conditionally heteroskedastic models: theory with applications to the term structure of interest rates', *Economic Notes*, **2**, 327–52.
- Glosten, L., R. Jagannathan and D. Runkle (1993), 'On the relation between the expected value and the volatility on the nominal excess returns on stocks', *Journal of Finance*, **48**, 1779–1801.
- Gouriéroux, C. and A. Monfort (1992), 'Qualitative threshold ARCH models', *Journal of Econometrics*, **52**, 159–99.
- Granger, C. and T. Teräsvirta (1993), *Modeling Nonlinear Economic Relationships*, Oxford University Press, Oxford.
- Harvey, A., E. Ruiz and N. Shephard (1994), 'Multivariate stochastic variance models', *Review of Economic Studies*, **61**, 247–64.
- Harvey, A. and N. Shephard (1993a), 'The econometrics of stochastic volatility', London School of Economics, Financial Markets Group, Discussion Paper 166.
- Harvey, A. and N. Shephard (1993b), 'Estimation and testing of stochastic variance models', STICERD Econometrics Discussion Paper, London School of Economics.
- Hull, J. and A. White (1987), 'The pricing of options on assets with stochastic volatilities', *Journal of Finance*, **42**, 281–300.
- Longstaff, F. and E. Schwartz (1992), 'Interest-rate volatility and the term structure: a two factor general equilibrium model', *Journal of Finance*, **47**, 1259–82.
- Mandelbrot, B. (1963), 'The variation of certain speculative prices', *Journal of Business*, **36**, 394–419.
- Melino, A. and S. Turnbull (1990), 'Pricing foreign currency options with stochastic volatility', *Journal of Econometrics*, **45**, 239–66.

- Modigliani, F. and M. Miller (1958), 'The cost of capital, corporation finance and the theory of investment', *American Economic Review*, **48**, 261–97.
- Nelson, D. (1990), 'ARCH models as diffusion approximations', *Journal of Econometrics*, **45**, 7–38.
- Nelson, D. (1991), 'Conditional heteroskedasticity in asset returns: a new approach', *Econometrica*, **59**, 347–70.
- Nelson, D. and D. Foster (1994), 'Asymptotic filtering theory for univariate ARCH models', *Econometrica*, **62**, 1–41.
- Pagan, A. and G. Schwert (1990), 'Alternative models for conditional stock volatility', *Journal of Econometrics*, **45**, 267–90.
- Papoulis, A. (1965), *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York.
- Rabemananjara, R. and J. M. Zakoïan (1993), 'Threshold ARCH models and asymmetries in volatility', *Journal of Applied Econometrics*, **8**, 31–49.
- Schwarz, G. (1978), 'Estimating the dimension of a model', *Annals of Statistics*, **6**, 461–4.
- Wiggins, J. B. (1987), 'Option values under stochastic volatility: theory and empirical estimates', *Journal of Financial Economics*, **19**, 351–72.
- Wong, E. (1971), *Stochastic Processes in Information and Dynamical Systems*, McGraw-Hill, New York.
- Zakoïan, J. M. (1994), 'Threshold heteroskedastic models', *Journal of Economic Dynamics and Control*, **18**, 931–55.