Closing the GARCH gap:
Continuous time GARCH modeling

Feike C. Drost, Bas J.M. Werker*

CentER and Department of Econometrics, Tilburg University, 5000 LE Tilburg, The Netherlands

Abstract

It is the purpose of this paper to build a bridge between continuous time models, which are central in the modern finance literature, and (weak) GARCH processes in discrete time, which often provide parsimonious descriptions of the observed data. The properties of continuous time processes which exhibit GARCH-type behavior at all discrete frequencies will be discussed. Several examples of such processes illustrate the general theory. The class of continuous time GARCH models can be divided into two subclasses. In the first group (GARCH diffusions) the sample paths are smooth and in the other group (GARCH jump-diffusions) the sample paths are erratic. A simple, complete characterization of both types is given in terms of the kurtosis of the *observed* discrete time data. These two groups of GARCH processes can be described by three and four coefficients, respectively. Explicit formulas of all implied discrete time weak GARCH parameters are available. Moreover, knowledge of the discrete time GARCH parameters at only one frequency completely determines the continuous time coefficients of the GARCH process. So, in estimating a continuous time GARCH process it suffices to estimate the discrete time GARCH parameters for the available data frequency. The analysis carries over to models with an autoregressive component.

Key words: GARCH; Diffusions; Jump-diffusions; Continuous time modeling
JEL classification: C22

1. Introduction

Since the seminal work of Black and Scholes (1973) continuous time models are one of the major tools in theoretical financial economics. They are used

*Corresponding author.

in general asset pricing theory (see, e.g., Huang, 1987; Cox, Ingersoll, and Ross, 1985a,b) and, more specific, in option pricing theory (see, e.g., Johnson and Shanno, 1987; Scott, 1987; Melino and Turnbull, 1990; Amin and Ng, 1993). These recent papers allow explicitly for a state variable influencing the asset price. Especially in the option pricing papers the volatility of the price process attracted much attention as an unobserved state variable. Pricing models for derivative securities heavily depend on the underlying model in continuous time (see, e.g., Melino and Turnbull, 1990). Usually, the validity of these continuous time models is not straightforward to check because data are available at discrete time only. In the empirical literature it is well-known that GARCH(1,1) processes often yield parsimonious representations of the observed data at almost every frequency. It is natural to ask whether continuous time models can be compatible with discrete time GARCH(1,1) pro- cesses at every (discrete) frequency. It turns out that this class of continuous time GARCH processes is rich enough to contain both diffusions and jump processes. While recent literature uses, almost without exception, discrete time models to approximate models in continuous time (see, e.g., Gourieroux, Monfort, and Renault, 1992; Nelson and Foster, 1994), this paper derives exact properties of the underlying continuous time GARCH process. Several examples are given.

We derive a simple criterion to discriminate between the smooth subgroup of continuous time GARCH models and the subgroup containing jumps. We show that it is sufficient to know the kurtosis of the implied discrete time difference process (at an arbitrary frequency) to distinguish between GARCH diffusions and jump-diffusions. Recognition of jumps is important in valuing derivative securi- ties. Out-of-the-money call options close to maturity will be virtually worthless if the underlying price process follows a diffusion while they will be valuable if the price process exhibits jumps. Diffusion models will underprice these options. For that reason Jorion (1988) has proposed a test procedure for the presence of jumps which relies on the normality of the conditional distribution of the nonjump component.

Moreover, we show that the assumption of an underlying continuous time GARCH model leads to kurtosis parameters of the corresponding discrete time processes which are necessarily strictly larger than three, implying heavy tails. This confirms the results of Drost and Nijman (1993). They observe that not every discrete time GARCH process can arise as the sum of underlying higher- frequency GARCH processes. Many authors explicitly introduce heavy-tailed in- novation distributions, such as student t-distributions, to capture this phenomenon. In fact, we show that the common, implicit assumption of an underlying model in continuous time already implies the appearance of heavy tails. Normal inno- vations are excluded at any frequency. This is in line with the empirical finding that conditional distributions are leptokurtic (see, e.g., Diebold, 1988; Bollerslev, Chou, and Kroner, 1992). This also implies that Jorion's (1988) testing procedure

(directed to normal innovations) should be adapted to include other distributions. This is discussed in more detail in Drost, Nijman, and Werker (1994).

Finally, we show that the coefficients of a continuous time GARCH process can be identified from the discrete time weak GARCH parameters at any arbitrary frequency and vice versa. This relation can be used to get fast, simple, consistent, correlation-based estimators of the parameters in the underlying continuous time model (see Drost and Nijman, 1993). In this way one may avoid the use of the recently developed simulation based estimators. These latter simulation methods are developed to estimate quite general models in continuous time (Duffie and Singleton, 1993; Gallant and Tauchen, 1992; Gourieroux, Monfort, and Renault, 1992). Of course, the efficiency of these complicated, time-consuming methods is likely to be higher than correlation-based methods since the criterion function is close to the true maximum likelihood equations. On the other hand, however, extra bias terms are introduced by the discrete time approximations of the underlying continuous time model. Probably, other commonly used estimators, like quasi maximum likelihood and semi-parametric procedures (see, e.g., Weiss, 1986; Linton, 1993; Drost and Klaassen, 1996), are also consistent (see, e.g., the small-scale simulation study of Drost and Nijman, 1993).

The paper is organized along the following lines. In Section 2 the concept of continuous time GARCH processes is introduced and illustrated by some examples. It will be shown that this class can be divided into two subgroups. In one of the groups we have smooth sample paths. These processes are called GARCH diffusions and are discussed in Section 3. Section 4 is devoted to the other group: GARCH jump-diffusions. For both subclasses the process will be characterized by a parameter vector of dimension three and four, respectively. These coefficients completely determine the discrete time weak GARCH parameters at all frequencies. A large variety of examples is included for both groups and an empirical example illustrates the general theory. The analysis carries over to models in which an autoregressive component is included (Section 5). Finally in Section 6 we will discuss some more implications of our results and conclude.

2. Continuous time GARCH processes

This section introduces the class of continuous time processes which exhibit GARCH-type behavior at all discrete frequencies. To make explicit calculations possible we concentrate on GARCH(1,1) processes. Of course, the theoretical framework of continuous time processes with GARCH behavior is easily extended to the general case. In this general setting, however, parameter restrictions and explicit formulas need numerical procedures and cannot be given in a closed form as in the GARCH(1,1) case. Restricting attention to GARCH(1,1) in the remainder of the paper we will simplify notation by deleting the orders and writing GARCH.

It seems natural to call a continuous time process $\{Y_t, t \geqslant 0\}$ GARCH if the first differences of the implied discrete time processes $\{Y_t, t \in h\mathbb{N}\}$ are GARCH for all fixed $h > 0$. Generally, however, $\{Y_{t+h} - Y_t, t \in h\mathbb{N}\}$ cannot be GARCH in the sense of Engle (1982) and Bollerslev (1986) for every $h > 0$ since Drost and Nijman (1993) have shown that this classical class is not closed under temporal aggregation. Instead of requiring that these differences are GARCH in this strong sense we will rely upon a weak GARCH definition (see Drost and Nijman, 1993).

Definition 2.1. Suppose $h > 0$. A symmetric discrete time process $\{y_{(h)t}, t \in h\mathbb{N}\}$ with finite fourth moments is called weak GARCH with parameter $\zeta_h = (\psi_h, \alpha_h, \beta_h, \kappa_h)$ if there exists a covariance-stationary process $\{\sigma^2_{(h)t}, t \in h\mathbb{N}\}$ with

$$\sigma^2_{(h)t+h} = \psi_h + \alpha_h y^2_{(h)t} + \beta_h \sigma^2_{(h)t}, \qquad t \in h\mathbb{N}, \tag{2.1}$$

such that, for $t \in h\mathbb{N}$, $\sigma^2_{(h)t}$ is the best linear predictor of $y^2_{(h)t}$ in terms of 1, $\sigma^2_{(h)0}$, and lagged values of $y_{(h)t}$ and $y^2_{(h)t}$. The parameter $\kappa_h = \mathrm{E}y^4_{(h)t}/(\mathrm{E}y^2_{(h)t})^2$ denotes the kurtosis of the process.

Throughout we assume the usual parameter restrictions $\psi_h > 0$, $\alpha_h \geqslant 0$, and either $\alpha_h = 0$ (and thus $\beta_h = 0$ for identifiability reasons) or $0 < \alpha_h + \beta_h < 1$. It is easy to see that the usual definition of GARCH with symmetric innovations and finite fourth moments (which will be called strong GARCH from now on) implies the weak GARCH one. In the general definition we still have $\mathrm{E}y^2_{(h)t} = \psi_h/(1 - \alpha_h - \beta_h)$. For reference we define the pseudo-kurtosis of the rescaled residuals $\xi_{(h)t} = y_{(h)t}/\sigma_{(h)t}$ by

$$\tilde{\kappa}_h = \kappa_h \frac{1 - (\alpha_h + \beta_h)^2 + \alpha^2_h}{1 - (\alpha_h + \beta_h)^2 + \alpha^2_h \kappa_h}. \tag{2.2}$$

This pseudo-kurtosis of rescaled residuals is the kurtosis of the innovations if the process is strong GARCH.

The class of weak GARCH models is closed under temporal aggregation (see Appendix A). Therefore, we adopt this weak definition in continuous time processes with conditional heteroskedastic behavior.

Definition 2.2 (GARCH process). A continuous time process $\{Y_t, t \geqslant 0\}$ is called GARCH if, for each starting time t_0 and each fixed time interval $h > 0$, the implied discrete time process $\{Y_{t_0+t+h} - Y_{t_0+t}, t \in h\mathbb{N}\}$ is weak GARCH with parameter vector $\zeta_h = (\psi_h, \alpha_h, \beta_h, \kappa_h)$.

Before deriving the implications of Definition 2.2 we will give four trivial examples. General classes of continuous time GARCH processes will be discussed in Sections 3–5.

Example 2.1. The most simple example is Brownian motion with variance parameter σ^2. At frequency h it satisfies the definition with $\zeta_h = (h\sigma^2, 0, 0, 3)$.

Example 2.2. Another simple example is the compound Poisson process, where the inter-arrival periods between jumps are assumed to be i.i.d. drawings from an exponential distribution with mean inter-arrival time μ and where the jumps are realizations of independent normals with variance σ^2. To verify the conditions let N_t be the number of jumps until time t and let X_i be the ith jump. Then $\{Y_{t+h} - Y_t, t \in h\mathbb{N}\}$ is an i.i.d. sequence of random variables with the same distribution as $Y_h = Y_h - Y_0 = \sum_{i=1}^{N_h} X_i$. Hence, using $EY_h = EY_h^3 = 0$, $EY_h^2 = \sigma^2(h/\mu)$, and $EY_h^4 = 3\sigma^4(h/\mu)^2 + 3\sigma^4(h/\mu)$, $\{Y_{t+h} - Y_t, t \in h\mathbb{N}\}$ is weak GARCH with $\zeta_h = (\sigma^2(h/\mu), \ 0, \ 0, \ 3 + 3\mu/h)$.

Example 2.3. A less simple example with continuous sample paths is the diffusion process given by the following system of differential equations:

$$\mathrm{d}Y_t = \sigma_t \, \mathrm{d}W_{(1)t}, \tag{2.3}$$

$$\mathrm{d}\sigma_t^2 = \theta(\omega - \sigma_t^2)\,\mathrm{d}t + \sqrt{2\lambda\theta}\,\sigma_t^2\,\mathrm{d}W_{(2)t}, \tag{2.4}$$

where $W_{(1)t}$ and $W_{(2)t}$ are independent Brownian motions, $\omega > 0$, $\theta > 0$, and $\lambda \in (0,1)$. [Nelson (1990) considers a slightly more general system with less parameter restrictions but we need the existence of fourth moments to be able to apply the aggregation results of Appendix A.] Nelson (1990) shows that these equations can be approximated by a sequence of discrete time GARCH processes with i.i.d. normal innovations. The hth approximating process is defined on the time scale $h\mathbb{N}$. Of course, all aggregates of every element in this approximating sequence are weak GARCH by Theorem A.1. This suggests that the corresponding limiting continuous time process for $h \downarrow 0$ is likely to be a GARCH process. A formal proof using stochastic calculus is given in Appendix C for a much more general class of GARCH processes, see also Sections 3 and 4. The vector ζ_h is given in Proposition 3.1.

Example 2.4. From the diffusion (2.3)–(2.4) one can easily construct a GARCH process with jumps by, e.g., adding an independent compound Poisson process to the solution of these differential equations, see Example 4.2. Compare Merton (1990, Sec. 9.2) and Amin (1993).

Theorem A.1 (see also Drost and Nijman, 1993; Example 2) induces several relationships between the parameters at different frequencies. Since we are working in a continuous time framework, we have an infinite number of equations. One might expect that there are four free parameters in a GARCH process. However, it will be shown below that, under the assumption of smooth sample paths, one only has three free parameters [for example, the diffusion (2.3)–(2.4)]. For nonsmooth GARCH processes there are still four parameters. So, the assumption

of an underlying diffusion in continuous time reduces the number of free parameters by one; delete, e.g., the kurtosis parameter κ_h in the definition of Drost and Nijman (1993) since it will be completely determined by the variance parameters α_h and β_h. This will have important implications in the sequel. To obtain these results we need the following regularity assumption.

Assumption A. The vector ζ_h is a continuous function in h.

Without this assumption Appendix A already implies that, for each fixed $h_0 > 0$, the parameter function $\zeta : h_0 \mathbb{Q}^+ \to \mathbb{R}^4$ is continuous and, hence, $\zeta(\cdot)$ is smooth on dense subsets of $\mathbb{R}^+$. Assumption A only excludes the possibility of a completely different behavior of the parameter vector on mutually exclusive dense subsets of $\mathbb{R}^+$. Therefore, Assumption A is harmless.

Our first result shows that the class of GARCH processes can be divided into two groups. These groups are distinguished by the behavior of $E|Y_h - Y_0|^4$.

Theorem 2.1. Let $\{Y_t, t \geq 0\}$ be a GARCH process and assume that Assumption A is fulfilled. Then $f(h) = E|Y_{t+h} - Y_t|^4$ is a continuous function of h not depending on $t \geq 0$, $f(h)/h^2$ converges in $(0, \infty)$ as $h \to \infty$, and either $f(h)/h^2$ or $f(h)/h$ converges in $(0, \infty)$ as $h \downarrow 0$.

Proof. See Appendix B.   □

The behavior of $f(h)$ is an important tool to characterize the level of smoothness of a continuous time process. If $f(h)/h^2$ is bounded, Kolmogorov's criterion (see, e.g., Theorem I.1.8 in Revuz and Yor, 1991) implies that $\{Y_t\}$ has a modification[1] with continuous sample paths. This group of GARCH processes will be called GARCH diffusions. The other group is not as smooth as, e.g., Brownian motion since the fourth moments are only of the order h, similar to a compound Poisson process. Therefore, these processes are called GARCH jump-diffusions.

Definition 2.3 (GARCH diffusion/jump-diffusion). Let $\{Y_t, t \geq 0\}$ be a continuous time GARCH process such that Assumption A holds. If $f(h)/h^2$ is bounded, then $\{Y_t, t \geq 0\}$ is called a GARCH diffusion. Otherwise it is called a GARCH jump-diffusion.

3. GARCH diffusions

In this section we derive some general results for GARCH diffusions. These results are exemplified by a broad class of GARCH diffusions (including the ones given in Section 2) and are applied to exchange rates.

---

[1] A modification $\{\tilde{Y}_t, t \geq 0\}$ of $\{Y_t, t \geq 0\}$ is a process satisfying $\tilde{Y}_t = Y_t$ (a.s.) for all $t \geq 0$.

The next proposition shows that GARCH diffusions can in fact be characterized by three coefficients, say $\omega > 0$, $\theta > 0$, and $\lambda \in (0,1)$. The parameter ζ_h is uniquely determined by these coefficients at every frequency h, and vice versa. Therefore, we will call ω, θ, and λ the characterizing coefficients of a GARCH diffusion. As an immediate consequence of the proposition we obtain information about the rates at which $\zeta_h \to (0, 0, 1, 3/(1 - \lambda))$. This is summarized in a corollary. These rates are in agreement with the rates Nelson (1990) uses to obtain the system of differential equations (2.3)–(2.4). We have parametrized these equations such that the parameters ω, θ, and λ are just the coefficients of the corresponding GARCH diffusion.

Proposition 3.1. *Let $\{Y_t, t \geqslant 0\}$ be a GARCH diffusion with parameter vectors $\zeta_h = (\psi_h, \alpha_h, \beta_h, \kappa_h)$ and suppose $\alpha_{h_0} \in (0,1)$ for some $h_0 > 0$. Then there exist $\omega \in (0, \infty)$, $\theta \in (0, \infty)$, $\lambda \in (0,1)$, and c_h given by*

$$c_h = \frac{4\{\exp(-h\theta) - 1 + h\theta\} + 2h\theta\{1 + h\theta(1 - \lambda)/\lambda\}}{1 - \exp(-2h\theta)}, \tag{3.1}$$

such that ζ_h (with $|\beta_h| < 1$) is determined by

$$\psi_h = h\omega\{1 - \exp(-h\theta)\}, \quad \kappa_h = 3 + 6\frac{\lambda}{1 - \lambda}\frac{\exp(-h\theta) - 1 + h\theta}{(h\theta)^2}, \tag{3.2}$$

$$\alpha_h = \exp(-h\theta) - \beta_h, \quad \frac{\beta_h}{1 + \beta_h^2} = \frac{c_h \exp(-h\theta) - 1}{c_h\{1 + \exp(-2h\theta)\} - 2}. \tag{3.3}$$

Proof. See Appendix B. □

Corollary 3.2. *Let $\{Y_t, t \geqslant 0\}$ be a GARCH diffusion with coefficients ω, θ, and λ. Then, as $h \downarrow 0$,*

$$\frac{\psi_h/h}{1 - \alpha_h - \beta_h} \to \omega, \quad \frac{1 - \alpha_h - \beta_h}{h} \to \theta,$$

$$\frac{\alpha_h^2}{1 - \alpha_h - \beta_h} \to \lambda, \quad \kappa_h \to 3/(1 - \lambda). \tag{3.4}$$

Moreover, $\mathrm{E}|Y_h - Y_0|^4/h^2 \to 3\omega^2/(1 - \lambda)$ as $h \downarrow 0$ and $\mathrm{E}|Y_h - Y_0|^4/h^2 \to 3\omega^2$ as $h \to \infty$.

Remark 3.1. If $\alpha_h = 0$ for all $h > 0$, then also $\beta_h = 0$ for all $h > 0$ (otherwise the parameters are not identifiable). Such continuous white noise processes, like Brownian motions, are obtained as limits in Proposition 3.1 with $\omega \in (0, \infty)$, $\theta = \infty$, $\lambda = 0$, and $\zeta_h = (h\omega, 0, 0, 3)$. As an example Eqs. (2.3) and (2.4) are still valid with solutions $Y_t = \sqrt{\omega}\, W_{(1)t}$ and $\sigma_t^2 = \omega$ (Brownian Motion).
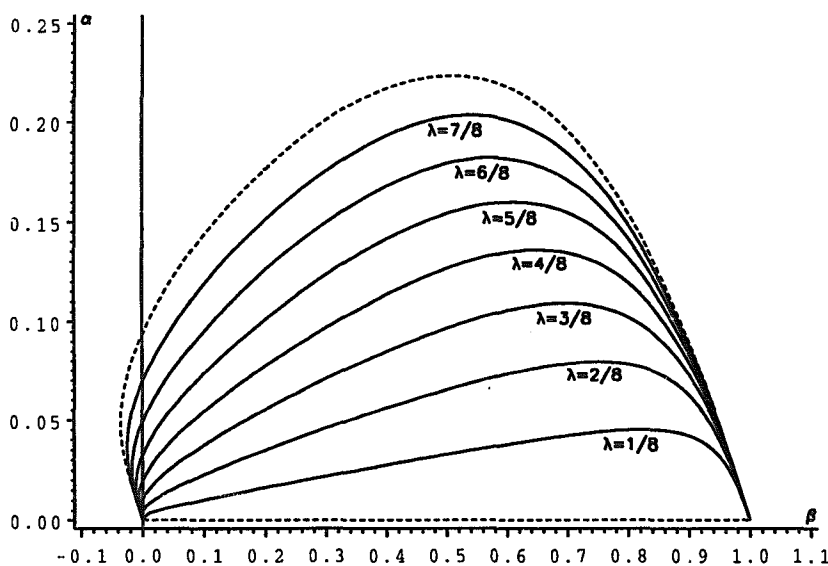
Fig. 3.1. Discrete time variance parameters of GARCH diffusions for several values of the coefficient λ.

The quadratic equation in β_h in (3.3) always admits exactly one solution with $|\beta_h| < 1$. Observe that β_h will become negative for very large values of h. At first sight, this seems to violate the parameter constraints of Nelson and Cao (1992), however, they are fine as coefficients in the linear projections underlying the weak GARCH formulations. Note that ω is a scale parameter. Since θ only appears in the form $h\theta$ it is normalized by the choice of the time unit. The parameter λ determines the slope of the kurtosis κ_h, see also the discussion below and Fig. 3.1.

Proposition 3.1 has several important implications. First, note that three of the four components of ζ_h, say ψ_h, α_h, and β_h, at some given fixed frequency h determine the coefficients ω, θ, and λ and, hence, they also fix the kurtosis κ_h. Since ω is merely a scale parameter and θ is a normalizing constant with respect to the time unit we concentrate on the parameter λ. Straightforward calculations show

$$\lambda = \frac{2\ln^2(\alpha_h + \beta_h)}{\frac{\{1-(\alpha_h+\beta_h)^2\}(1-\beta_h)^2}{\alpha_h\{1-\beta_h(\alpha_h+\beta_h)\}} + 6\ln(\alpha_h + \beta_h) + 2\ln^2(\alpha_h + \beta_h) + 4(1 - \alpha_h - \beta_h)}.$$

Observe that the right-hand side will not depend on h. So, the variance parameters at one frequency also uniquely determine the variance parameters at all other frequencies. This is illustrated in Fig. 3.1, where the lines correspond to GARCH diffusions with different values of the slope parameter λ. The points at some given line are the variance parameters α_h and β_h of the discrete time weak GARCH

processes associated to the GARCH diff: ion. High-frequency parameters are close to $\alpha = 0$ and $\beta = 1$. Moving along a line to the left corresponds to lower (and lower) sample frequencies. This figure is comparable to Fig. 2 in Drost and Nijman (1993). The main difference is that, in our situation, we do not have different lines passing through one point. This is caused by the assumption of an underlying diffusion in continuous time, implying that the kurtosis κ_h is completely fixed by α_h and β_h. In GARCH diffusions the kurtosis will not vary freely like in Drost and Nijman (1993). See, however, Section 4 for GARCH jump-diffusions.

Secondly, we direct attention to the kurtosis value of the process. It is clear from (3.2) that the kurtosis of the discrete time weak GARCH processes corresponding to a GARCH diffusion is strictly larger than three. This corresponds to the stylized fact that financial data have fat tails. The definition of a GARCH diffusion immediately yields this property and, therefore, these processes seem to be useful while modeling financial data. In classical analyses of GARCH processes, where the rescaled innovations are assumed to be independent, one also pays attention to the distributional aspects of the innovations. For strong GARCH processes the relationship between the kurtosis $\tilde{\kappa}_h$ of the innovations, the kurtosis κ_h of the GARCH process, and the GARCH parameters α_h and β_h is given by (2.2). In weak GARCH processes this parameter $\tilde{\kappa}_h$ is called the pseudo-kurtosis since the innovations are not i.i.d. (cf. Drost and Nijman, 1993). The formulas from Proposition 3.1 are substituted into the right-hand side of (2.2) to investigate whether GARCH diffusions imply leptokurtosis of the innovations, too. The pseudo-kurtosis is completely determined by the GARCH parameters α_h and β_h and one may verify that it is always larger than three, suggesting heavy-tailed innovations. Hence, the existence of an underlying diffusion in a conditional heteroskedastic framework confirms the empirical evidence that innovations are heavy-tailed; see Diebold (1988). This is outlined in Fig. 3.2. Contour lines are given for the pseudo-kurtosis in the area of the (α_h, β_h) space where GARCH diffusions are applicable.

We present two additional examples. Example 3.1 introduces a general class of GARCH diffusions and Example 3.2 discusses contemporaneous aggregation of GARCH diffusions.

Example 3.1. Let $\{W_t, t \geqslant 0\}$ be a standardized Brownian motion, $EW_t^2 = t$, independent of the standardized Lévy process $\{L_t, t \geqslant 0\}$, $EL_t^2 = t$. Then the solution $\{Y_t, t \geqslant 0\}$ of the system of differential equations

$$dY_t = \sigma_{t-} \, dW_t,$$ \hfill (3.5)

$$d\sigma_t^2 = \theta(\omega - \sigma_{t-}^2) \, dt + \sqrt{2\lambda\theta} \, \sigma_{t-}^2 \, dL_t,$$ \hfill (3.6)

with $\omega > 0$, $\theta > 0$, and $\lambda \in (0,1)$, is a GARCH diffusion with characterizing coefficients ω, θ, and λ.
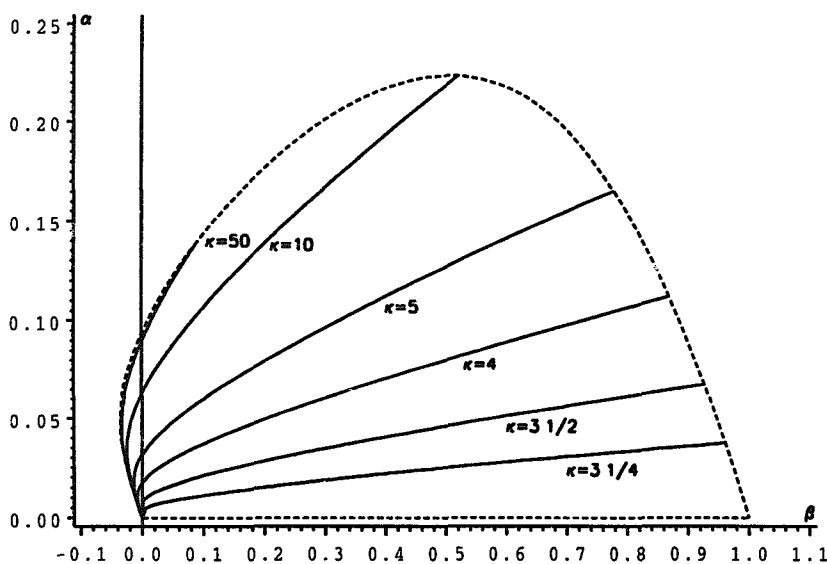
Fig. 3.2. Pseudo-kurtosis contour lines of GARCH diffusions.

Proof. See Appendix C. □

Recall that Lévy processes have independent stationary increments and that these processes will exhibit jumps unless $\{L_t\}$ is a Brownian motion (see Theorem II.38 of Protter, 1990). Important examples are the compound Poisson process and the Gamma process (see, e.g., Heston, 1993). Several special cases of the class of processes defined by (3.5) and (3.6) have been studied before. If $\sigma_t^2 = \omega$ is constant, then the implied spot price $S_t = \exp(Y_t)$ is a geometric Brownian motion. The system of differential equations (2.3)–(2.4) (Nelson, 1990) is another example of (3.5) and (3.6) where the Lévy process is specialized to Brownian motion. Eq. (3.6) explicitly allows for volatility processes $\{\sigma_t^2\}$ with jumps by taking other Lévy processes. Naik (1993) discusses the pricing of options when the volatility process exhibits jumps. Note that the characterizing coefficients do not depend on the choice of $\{L_t\}$. So, the same parameter configuration holds for all solutions of (3.5) and (3.6). The distribution of the continuous time process $\{Y_t\}$ is not completely determined by the coefficients ω, θ, and λ but also by the choice of $\{L_t\}$. This implies that estimation of (3.5)–(3.6) via GARCH parameters does not depend on the specification of $\{L_t\}$.

Example 3.2. Let $\{Y_{(i)t}, t \geq 0\}$, $i = 1, \ldots, k$, be independent GARCH diffusions with characterizing coefficients $(\omega_i, \theta_i, \lambda_i)$ with either $\theta_i = \theta_0 \in (0, \infty)$ or jointly $\theta_i = \infty$ and $\lambda_i = 0$. So the ith GARCH diffusion is either a stochastic process satisfying the conditions of Proposition 3.1 or a continuous white noise process as sketched in Remark 3.1. Suppose, for simplicity, that at least one of the θ_i's

equals θ_0. Then the sum process $\{Y_t = \sum_{i=1}^k Y_{(i)t}, t \geqslant 0\}$ is a GARCH diffusion with coefficients $\omega = \sum_{i=1}^k \omega_i$, $\theta = \theta_0$, and λ determined by

$$\frac{\lambda}{1-\lambda} = \sum_{i=1}^k \frac{\lambda_i}{1-\lambda_i}\omega_i^2 \bigg/ \left(\sum_{i=1}^k \omega_i\right)^2. \tag{3.7}$$

Proof. Along the lines of Nijman and Sentana (1993) one obtains that the sum process is GARCH at each discrete frequency with, e.g., $\psi_h = \sum_{i=1}^k \omega_i h\{1 - \exp(-\theta_0 h)\}$ and $\alpha_h + \beta_h = \exp(-\theta_0 h)$. Obviously the sum process is a GARCH diffusion. The relations concerning ψ_h and $\alpha_h + \beta_h$ determine ω and θ. To obtain the required equation for λ, observe that the relationship between the kurtosis κ_h of the sum process and the kurtosises $\kappa_{(i)h}$ of the separate parts is given by

$$\kappa_h - 3 = \sum_{i=1}^k (\kappa_{(i)h} - 3)\omega_i^2 \bigg/ \left(\sum_{i=1}^k \omega_i\right)^2.$$

Using the GARCH diffusion property (3.2), one obtains (3.7) by taking the limit for $h \downarrow 0$. □

Aggregation of a large set of independent 'balanced' GARCH diffusions yields a GARCH diffusion with a λ value close to zero, implying $\kappa_h \approx 3$, for all $h > 0$. As usual, aggregated data exhibits less leptokurtosis.

As a special case of the formulas above we obtain that the sum of the GARCH diffusion (3.5)–(3.6) and a Brownian motion with variance parameter σ^2 is a GARCH diffusion with coefficients $\tilde{\omega} = \omega + \sigma^2$, $\tilde{\theta} = \theta$, and $\tilde{\lambda}$ determined by

$$\frac{\tilde{\lambda}}{1-\tilde{\lambda}} = \frac{\lambda}{1-\lambda}\omega^2/(\omega + \sigma^2)^2.$$

We conclude this section with an empirical example considering six exchange rates under the assumption that the underlying DGP is a GARCH diffusion in continuous time. The implications of jumps will be examined in Section 4. Our estimates of the characterizing coefficients θ and λ are obtained from Proposition 3.1 by plugging in the daily estimates of the GARCH parameters α and β as reported in Baillie and Bollerslev (1989). Their estimated value of the kurtosis is ignored in these calculations and it is confronted with the kurtosis implied by the assumed underlying GARCH diffusion. These results are presented in Table 3.1. For the JY/\$, FF/\$, and BP/\$ exchange rates the difference between the kurtosis implied by the GARCH process and the direct estimate is rather large. This suggests that the assumption of an underlying diffusion is not very realistic in these cases. Probably jumps are present. For the other exchange rates this difference is rather small and one may expect that a diffusion model yields a satisfactory description. Plugging in the estimates of the GARCH coefficients into, e.g., (3.5) and (3.6) yields an estimate of the DGP in continuous time. This estimate of the price process can be used to value options or to construct hedge portfolios.

Table 3.1
Estimates of GARCH parameters for six exchange rates; March 1, 1980 to January 28, 1985
The first three columns are direct daily GARCH estimates (Baillie and Bollerslev, 1989). The other columns contain the implied GARCH diffusion coefficients and the implied kurtosis.

	GARCH estimates			Diffusion estimates		
	α	β	$\tilde{\kappa}$	θ	λ	$\tilde{\kappa}$
JY/$	0.049	0.941	5.62	0.010	0.357	3.34
FF/$	0.114	0.829	4.92	0.059	0.533	4.07
BP/$	0.061	0.910	4.16	0.029	0.272	3.45
IL/$	0.113	0.848	3.89	0.040	0.637	4.04
GM/$	0.085	0.881	3.41	0.035	0.427	3.69
SF/$	0.073	0.907	3.41	0.020	0.450	3.56

4. GARCH jump-diffusions

This section contains the counterpart of Section 3: GARCH jump-diffusions. Similar results are derived for this nonsmooth subclass of continuous time GARCH models. The following proposition shows that these processes can be characterized by four coefficients.

Proposition 4.1. *Let* $\{Y_t, t \geq 0\}$ *be a GARCH jump-diffusion with parameter vectors* $\zeta_h = (\psi_h, \alpha_h, \beta_h, \kappa_h)$ *and suppose* $\alpha_{h_0} > 0$ *for some* $h_0 > 0$. *Then there exist* $\omega \in (0, \infty)$, $\theta \in (0, \infty)$, $\phi \in (0, \infty)$, *and* $v \in (0, \infty)$, *such that* ζ_h *(with* $|\beta_h| < 1$*) is determined by Eqs. (3.2) and (3.3) with* c_h *and* κ_h *replaced by*

$$c_h = \frac{4\{\exp(-h\theta) - 1 + h\theta\} + 2h\theta \left\{1 + \frac{v + 2h\theta}{v\phi(2+\phi)}\right\}}{1 - \exp(-2h\theta)}, \tag{4.1}$$

$$\kappa_h = 3 + \frac{v}{h\theta} + 3v\phi(2 + \phi)\frac{\exp(-h\theta) - 1 + h\theta}{(h\theta)^2}. \tag{4.2}$$

Proof. See Appendix B. □

Corollary 4.2. *Let* $\{Y_t, t \geq 0\}$ *be a GARCH jump-diffusion with coefficients* ω, θ, ϕ, *and* v. *Then, as* $h \downarrow 0$,

$$\frac{\psi_h/h}{1 - \alpha_h - \beta_h} \to \omega, \qquad \frac{1 - \alpha_h - \beta_h}{h} \to \theta,$$

$$\frac{\alpha_h}{1 - \alpha_h - \beta_h} \to \phi, \qquad h\theta\kappa_h \to v. \tag{4.3}$$

Moreover, $E|Y_h - Y_0|^4/h \to v\omega^2/\theta$ *as* $h \downarrow 0$ *and* $E|Y_h - Y_0|^4/h^2 \to 3\omega^2$ *as* $h \to \infty$.

Remark 4.1. If $\alpha_h = \beta_h = 0$ for all $h > 0$, then GARCH jump-diffusions are obtained as limits with $\omega \in (0, \infty)$, $\theta = \infty$, $\phi = 0$, $v = \infty$, $v^* = \lim_{h \downarrow 0} h\kappa_h \in (0, \infty)$, and $\zeta_h = (h\omega, 0, 0, 3 + v^*/h)$. An example is given by the compound Poisson process of Section 2 with $\omega = \sigma^2/\mu$ and $v^* = 3\mu$.

The discussion of GARCH diffusions carries over to the class of jump-diffusions. As before the time unit normalizes θ and scale is denoted by ω. The parameters ϕ and v are slope parameters. Similar to λ in diffusions, ϕ will denote slopes in the (α_h, β_h) plane (compare Fig. 3.1) while v determines the slope of the kurtosis at very high frequencies. In contrast to the situation in diffusions and due to the additional free kurtosis parameter we have four characterizing coefficients. The overidentifying restriction in Section 3 is missing. Finally, note that given the weak GARCH parameters α_h and β_h we obtain identical values for θ and c_h in Propositions 3.1 and 4.1, respectively. One readily verifies that the kurtosis for GARCH jump-diffusions is larger than the one for GARCH diffusions. As in GARCH diffusions this confirms the empirical finding of heavy tails both in the innovations and the log-prices themselves. It also yields another interpretation of Fig. 3.2. Given the weak GARCH parameters α_h and β_h this figure determines lower bounds for the pseudo-kurtosis. If the true kurtosis is larger than or equal to the value obtained from the figure, then an underlying jump process or diffusion is possible, respectively. Otherwise an underlying process in continuous time does not exist. This also explains why Drost and Nijman (1993) could not determine the weak GARCH parameters at very high frequencies in some special occasions. In these situations the kurtosis value is too low.

We present two additional examples. Example 4.1 introduces a general class of jump-diffusions and Example 4.2 discusses contemporaneous aggregation of GARCH processes.

Example 4.1. Let $\{L_t, t \geqslant 0\}$ and $\{M_t, t \geqslant 0\}$ be two independent standardized Lévy processes, $EL_t^2 = EM_t^2 = t$, and suppose that $\{L_t\}$ is symmetric with $v_L^* = EL_1^4 - 3 < \infty$. Consider the system of differential equations

$$dY_t = \sigma_{t-} \, dL_t, \tag{4.4}$$

$$d\sigma_t^2 = \theta(\omega - \sigma_{t-}^2) \, dt + \sqrt{2\eta\theta} \, \sigma_{t-}^2 \, dM_t, \tag{4.5}$$

with $\omega > 0$, $\theta > 0$, and $\eta \in (0, 1)$. If $\{L_t\}$ is Brownian motion, then we are in the situation of Example 3.1. Otherwise, the solution $\{Y_t, t \geqslant 0\}$ of (4.4) and (4.5) is a GARCH jump-diffusion with characterizing coefficients ω, θ, $v = \theta v_L^*/(1 - \eta)$, and ϕ determined by

$$v\phi(\phi + 2) = 2\eta/(1 - \eta).$$

Proof. See Appendix C. $\square$

Specializing $\{L_t\}$ to the sum of a Brownian motion and an independent compound Poisson process and assuming that the variance process $\{\sigma_t^2\}$ is constant yields the price process considered in Merton (1990, Sec. 9.2) and Amin (1993). The same form of $\{L_t\}$, but with stochastic volatility driven by (4.5), is discussed in Drost, Nijman, and Werker (1994). Eq. (4.4) explicitly allows for nonconstant volatility processes in addition to jumps. Note that the characterizing coefficients only depend on $\{L_t\}$ through v_L^*.

Example 4.2. Let $\{Y_{(i)t}, t \geqslant 0\}$, $i = 1,\ldots,k_1 + k_2$, be independent GARCH processes, where the first k_1 processes are GARCH diffusions with characterizing coefficients $(\omega_i, \theta_i, \lambda_i)$ with either $\theta_i = \theta_0 \in (0,\infty)$ or jointly $\theta_i = \infty$ and $\lambda_i = 0$ and where the latter k_2 processes are GARCH jump-diffusions with characterizing coefficients $(\omega_i, \theta_i, \phi_i, v_i)$ with either $\theta_i = \theta_0$ or jointly $\theta_i = v_i = \infty$ and $\phi_i = 0$. So the separate stochastic processes satisfy either the conditions of Propositions 3.1 and 4.1 or the ones sketched in Remarks 3.1 and 4.1. Assume, for simplicity, that at least one of the θ_i's equals θ_0 and redefine v_i by $v_i = \theta_0 \lim_{h\downarrow 0} h \kappa_{(i)h} = \theta_0 v_i^*$ if $v_i = \infty$. Then the sum process $\{Y_t = \sum_{i=1}^{k_1+k_2} Y_{(i)t}, t \geqslant 0\}$ is a GARCH jump-diffusion with coefficients $\omega = \sum_{i=1}^{k_1+k_2} \omega_i$, $\theta = \theta_0$, and ϕ and v determined by

$$v = \sum_{i=k_1+1}^{k_1+k_2} v_i \omega_i^2 \Big/ \left(\sum_{i=1}^{k_1+k_2} \omega_i \right)^2,$$

$$v\phi(2 + \phi) = \left\{ \sum_{i=1}^{k_1} 2 \frac{\lambda_i}{1 - \lambda_i} \omega_i^2 + \sum_{i=k_1+1}^{k_1+k_2} v_i \phi_i (2 + \phi_i) \omega_i^2 \right\} \Big/ \left(\sum_{i=1}^{k_1+k_2} \omega_i \right)^2.$$

The proof is completely similar to the proof of Example 3.2. Evaluate the limit of $h(\kappa_h - 3)$ both for $h \downarrow 0$ and $h \to \infty$.

As in Example 3.2, leptokurtosis is less pronounced in aggregated series. The parameter v will generally decrease to 0 as $k_1 + k_2 \to \infty$ and, hence, $\kappa_h \approx 3$ unless h small.

As a special case of the formulas above we obtain that the sum of the GARCH diffusion (3.5)–(3.6) and the compound Poisson process of Section 2 yields a GARCH jump-diffusion with coefficients $\tilde\omega = \omega + \sigma^2/\mu$, $\tilde\theta = \theta$, and $\tilde\phi$ and $\tilde v$ determined by

$$\tilde v = \theta \kappa \sigma^4 \mu^{-1} / (\omega + \sigma^2/\mu)^2,$$

$$\tilde v \tilde\phi(\tilde\phi + 2) = 2 \frac{\lambda}{1 - \lambda} \omega^2 / (\omega + \sigma^2/\mu)^2.$$

Finally, we reconsider the empirical example about exchange rates. Many empirical studies suggest the presence of jumps in exchange rates; cf., e.g., Jorion (1988) and Vlaar and Palm (1993). Large jumps may be caused by re-alignments but frequent small jumps have also been observed. The characterizing GARCH jump coefficients θ, ϕ, and v are obtained from Proposition 4.1 by

Table 4.1
Estimates of GARCH parameters for six exchange rates; March 1, 1980 to January 28, 1985

The first three columns are direct daily GARCH estimates (Baillie and Bollerslev, 1989). The other columns contain the implied GARCH jump coefficients.

	GARCH estimates			Jump estimates				
	α	β	$\tilde{\kappa}$	θ	v	ϕ	η	v_L^*
JY/$	0.049	0.941	5.62	0.010	0.044	8.03	0.640	1.58
FF/$	0.114	0.829	4.92	0.059	0.060	6.65	0.634	0.38
BP/$	0.061	0.910	4.16	0.029	0.021	5.98	0.337	0.48
IL/$	0.113	0.848	3.89†					
GM/$	0.085	0.881	3.41†					
SF/$	0.073	0.907	3.41†					

† The estimated kurtosis value is too small to admit an underlying jump-diffusion in continuous time.

plugging in the daily estimates of the GARCH parameters α and β and the estimated kurtosis parameter $\tilde{\kappa}$ (Baillie and Bollerslev, 1989). The results are given in Table 4.1. For the IL/$, GM/$, and SF/$ exchange rates it is not possible to obtain the characterizing jump parameters since the estimates of the kurtosis are somewhat smaller than the corresponding critical values (obtainable from Fig. 3.2). Observe, however, that the difference between the estimated kurtosis and this value is not very large. Hence, diffusion models or jump-diffusions with less pronounced jump components seem to be a good descriptions in these cases. As suggested in Section 3, the figures for the other exchange rates point to more pronounced jumps. In Table 4.1 we also give the corresponding values of η and v_L^* in Example 4.1. In this way we obtain an estimate of the underlying DGP in continuous time. The value v_L^* fixes the kurtosis parameter of $\{L_t\}$. Other characteristics of the Lévy processes can be chosen freely by the researcher. Estimation and testing in GARCH (jump-)diffusions are investigated in Drost, Nijman, and Werker (1994).

5. Extension to autoregressive components

Some financial series, like, e.g., interest rates, exhibit both autocorrelation and conditional heteroskedasticity. A continuous time model that is able to generate both features (and a possible trend) is given by the following system of differential equations

$$\mathrm{d}Y_t = (\delta - \tau Y_{t-})\,\mathrm{d}t + \sigma_{t-}\,\mathrm{d}L_t, \tag{5.1}$$

$$\mathrm{d}\sigma_t^2 = \theta(\omega - \sigma_{t-}^2)\,\mathrm{d}t + \sqrt{2\lambda\theta}\,\sigma_{t-}^2\,\mathrm{d}M_t. \tag{5.2}$$

Note that $\delta = \tau = 0$ leads to the GARCH processes discussed in Examples 3.1 and 4.1 and that $\delta = \theta = 0$ leads to the familiar Ornstein–Uhlenbeck process.

Similar to the derivations in Sections 2–4 we may define an autoregressive GARCH process as a process which, for each discrete frequency, is a shifted auto-regressive time-series model of order one with GARCH innovations. By putting the autoregressive parameter $\exp(-h\tau)$ equal to one, the unit root case, we obtain Definition 2.2. Although we do not go into details one can derive results similar to the ones before, using a generalization of Theorem A.1 for autoregressive GARCH models. Explicit formulas of this generalization can be obtained along the lines in the proofs of Drost and Nijman (1993). Their formulas in Examples 1 and 2 are a special case if the autoregressive parameter is equal to zero and one, respectively. Theorem 2.1 also applies in this extended setting. Continuous time autoregressive GARCH models can be divided into a smooth and a nonsmooth class. The differential equation above generates examples in both groups. Similar to Propositions 3.1 and 4.1 the parameters of the implied discrete time models are determined by five and six coefficients, respectively. We have to add two parameters, say δ and τ, to account for the trend and the autoregressive component in the model. At frequency h the shift equals $\mu_h = \delta/\tau$ and the autoregressive parameter ρ_h is given by $\rho_h = \exp(-h\tau)$. Furthermore, the discrete time scale parameter is given by

$$\psi_h = h\omega\{1 - \exp(-h\theta)\}\frac{1 - \exp(-2h\tau)}{2h\tau}. \tag{5.3}$$

The GARCH variance parameters are determined by

$$\alpha_h = \exp(-h\theta) - \beta_h, \quad \frac{\beta_h}{1 + \beta_h^2} = \frac{a_h \exp(-h\theta) - 1}{a_h\{1 + \exp(-2h\theta)\} - b_h}, \tag{5.4}$$

where a_h and b_h are some complicated formulas; see Appendix D. If $\tau = 0$, $b_h = 2$ and a_h specializes to the c_h values in Propositions 3.1 and 4.1. Finally the kurtosis values of the GARCH component in GARCH diffusions and GARCH jump-diffusions are given by

$$\kappa_h = 3 + 6\frac{\lambda}{1 - \lambda}A(h, \theta, \tau), \tag{5.5}$$

$$\kappa_h = 3 + \frac{v}{h\theta}\frac{1 + \exp(-2h\tau)}{2}\frac{2h\tau}{1 - \exp(-2h\tau)} + 3v\phi(2 + \phi)A(h, \theta, \tau), \tag{5.6}$$

respectively, with

$$A(h, \theta, \tau)$$
$$= \left(\frac{2h\tau}{1 - \exp(-2h\tau)}\right)^2 \frac{\exp(-h\{\theta + 2\tau\}) - \exp(-4h\tau) + h\frac{1 - \exp(-4h\tau)}{4h\tau}(\theta - 2\tau)}{h^2(\theta + 2\tau)(\theta - 2\tau)}.$$

If $\tau = 0$, then we obtain the propositions in Sections 3 and 4 as a special example. If $\tau = 2\theta$, replace the expression $A(h, \theta, \tau)$ by the corresponding limit.

6. Conclusions

In this paper we have shown that the common assumption of an underlying model in continuous time can perfectly agree with the empirical finding of GARCH at all discrete frequencies. An explicit one–one relationship between parameters in continuous and discrete time models is available for the GARCH (1,1) case. A computer program evaluating these expressions is available on request from the authors. Moreover, these relations can be used for testing and fast estimation, avoiding simulation techniques. The class of continuous GARCH models contains models with continuous as well as jumpy sample paths. Our results suggest straightforward tests to distinguish between these two classes. Finally, our results provide an explanation why fat-tailed conditional distributions are obtained, without exception, in empirical work.

Appendix A: Discrete time GARCH aggregation

We introduce the following convention: an element x belongs to a set, like $h\mathbb{N}$ or $h\mathbb{Q}^+$, if x/h belongs to $\mathbb{N}$ or $\mathbb{Q}^+$, respectively. Drost and Nijman (1993, Example 2) shows that the class of weak GARCH processes is closed under temporal aggregation.

Theorem A.1. Let $h > 0$ and suppose $\{y_{(h)t}, t \in h\mathbb{N}\}$ is a weak GARCH process with parameter $\zeta_h = (\psi_h, \alpha_h, \beta_h, \kappa_h)$. Then, for each integer $m \geqslant 1$, the process $\{y_{(mh)t}^{(m)} = \sum_{i=0}^{m-1} y_{(h)t+ih}, t \in mh\mathbb{N}\}$ is symmetric weak GARCH with parameter $\zeta_{mh} = (\psi_{mh}, \alpha_{mh}, \beta_{mh}, \kappa_{mh})$ (with $|\beta_{mh}| < 1$) determined by

$$\psi_{mh} = m\psi_h \frac{1 - (\alpha_h + \beta_h)^m}{1 - (\alpha_h + \beta_h)}, \tag{A.1}$$

$$\alpha_{mh} = (\alpha_h + \beta_h)^m - \beta_{mh}, \tag{A.2}$$

$$\frac{\beta_{mh}}{1 + \beta_{mh}^2} = \frac{a(\alpha_h, \beta_h, \kappa_h, m)(\alpha_h + \beta_h)^m - b(\alpha_h, \beta_h, m)}{a(\alpha_h, \beta_h, \kappa_h, m)\{1 + (\alpha_h + \beta_h)^{2m}\} - 2b(\alpha_h, \beta_h, m)}, \tag{A.3}$$

$$\kappa_{mh} = 3 + \frac{\kappa_h - 3}{m} + 6(\kappa_h - 1)$$

$$\times \frac{\{m(1 - \alpha_h - \beta_h) - 1 + (\alpha_h + \beta_h)^m\}\{\alpha_h(1 - (\alpha_h + \beta_h)^2) + \alpha_h^2(\alpha_h + \beta_h)\}}{m^2(1 - \alpha_h - \beta_h)^2\{1 - (\alpha_h + \beta_h)^2 + \alpha_h^2\}}, \tag{A.4}$$

where

$$a(\alpha_h, \beta_h, \kappa_h, m)$$

$$= m(1 - \beta_h)^2 + 2m(m-1)\frac{(1 - \alpha_h - \beta_h)^2\{1 - (\alpha_h + \beta_h)^2 + \alpha_h^2\}}{(\kappa_h - 1)\{1 - (\alpha_h + \beta_h)^2\}}$$

$$+ 4\frac{\{m(1 - \alpha_h - \beta_h) - 1 + (\alpha_h + \beta_h)^m\}\{\alpha_h(1 - (\alpha_h + \beta_h)^2) + \alpha_h^2(\alpha_h + \beta_h)\}}{1 - (\alpha_h + \beta_h)^2},$$

$$b(\alpha_h, \beta_h, m) = \{\alpha_h(1 - (\alpha_h + \beta_h)^2) + \alpha_h^2(\alpha_h + \beta_h)\}\frac{1 - (\alpha_h + \beta_h)^{2m}}{1 - (\alpha_h + \beta_h)^2}.$$

Let q be the transfer function corresponding to Theorem A.1 that transforms high-frequency parameters into low-frequency ones, i.e., $q(\zeta_h, m) = \zeta_{mh}$. The interpretation of Theorem A.1 implies $q(q(\zeta_h, m), n) = q(\zeta_h, mn)$ for all integers m and n. Tedious calculations, using a formula handling package (e.g., Mathematica), show that the latter equality also holds true if the integers m and n are replaced by arbitrary reals. This observation will be useful in our derivations in a continuous time context. E.g., if a weak GARCH process with parameter ζ_h is known to be the aggregate over m periods of some other higher-frequency GARCH process, then the parameter of the latter high-frequency process is given by $\zeta_{h/m} = q(\zeta_h, 1/m)$. If one assumes that the observed process at frequency, say, g is infinitely divisible, i.e., if one assumes that for each integer m there exists an underlying high-frequency GARCH process such that the observed process is the sum over m periods of the high-frequency process, then the transfer function q determines the parameters by $\zeta_h = q(\zeta_g, h/g)$ for all $h \in g\mathbb{Q}^+$.

Appendix B: Proofs of main results

Proof of Theorem 2.1. Observe that the continuity of the GARCH parameters together with the remarks at the end of Appendix A imply that knowledge of the GARCH parameter at some specific frequency, say given ζ_g, completely determines ζ_h for all $h > 0$ by $\zeta_h = q(\zeta_g, h/g)$. Hence, if h_n is a sequence decreasing to zero as $n \to \infty$,

$$\zeta_h = q(\zeta_{h_n}, h/h_n) = \lim_{n \to \infty} q(\zeta_{h_n}, h/h_n). \tag{B.1}$$

Choose the sequence h_n such that

$$\frac{\psi_{h_n}/h_n}{1 - \alpha_{h_n} - \beta_{h_n}} \to \omega \in [0, \infty],$$

$$\frac{1 - \alpha_{h_n} - \beta_{h_n}}{h_n} \to \theta \in [0, \infty],$$

$$(1 - \alpha_{h_n} - \beta_{h_n})\kappa_{h_n} \to v \in [0, \infty],$$

$$(\kappa_{h_n} - 1)\frac{\alpha_{h_n}\{1 - (\alpha_{h_n} + \beta_{h_n})^2\} + \alpha_{h_n}^2(\alpha_{h_n} + \beta_{h_n})}{1 - (\alpha_{h_n} + \beta_{h_n})^2 + \alpha_{h_n}^2} \to \rho \in [0, \infty].$$

The calculations in the following equations are based on the combination of Theorem A.1 and (B.1).

$$\alpha_h + \beta_h = \lim_{n \to \infty} (\alpha_{h_n} + \beta_{h_n})^{h/h_n} = \exp(-h\theta).$$

Suppose $\alpha_{h_0} > 0$ for some $h_0 > 0$ (the case with $\alpha_h = \beta_h = 0$ for all $h > 0$ is simple). Since $0 < \alpha_{h_0} + \beta_{h_0} < 1$, this implies $\theta \in (0, \infty)$. Using $\theta < \infty$, one obtains in a similar manner

$$\psi_h = h\omega\{1 - \exp(-h\theta)\},$$

with $\omega \in (0, \infty)$, and

$$\kappa_h = 3 + \frac{v}{h\theta} + 6\rho\frac{\exp(-h\theta) - 1 + h\theta}{(h\theta)^2}. \tag{B.2}$$

By the weak GARCH assumption κ_h is finite for each h, implying $v \in [0, \infty)$ and $\rho \in [0, \infty)$. Observe that, for each frequency h, explicit formulas of ψ_h, $\alpha_h + \beta_h$, and κ_h are obtained only depending upon the limiting variables ω, θ, v, and ρ. This shows that these limits cannot depend upon the chosen sequence; the same values are obtained for all sequences tending to zero. The proof is completed by noting that $\lim_{h \to \infty} \mathrm{E}|Y_{t+h} - Y_t|^4/h^2 = 3\omega^2$ and

$$\lim_{h \downarrow 0} \mathrm{E}|Y_{t+h} - Y_t|^4/h^2 = \lim_{h \downarrow 0} \kappa_h \left(\frac{\psi_h/h}{1 - \alpha_h - \beta_h}\right)^2 = 3(1 + \rho)\omega^2$$

or

$$\lim_{h \downarrow 0} \mathrm{E}|Y_{t+h} - Y_t|^4/h = \lim_{h \downarrow 0} h\kappa_h \left(\frac{\psi_h/h}{1 - \alpha_h - \beta_h}\right)^2 = v\omega^2/\theta$$

if κ_h is bounded or unbounded near $h = 0$, respectively. $\square$

Proof of Proposition 3.1. We continue with the proof of Theorem 2.1 as starting point and consider the class of GARCH processes with bounded kurtosis, i.e., $v = 0$. Two of the required relations are already obtained. Suppose that, along the sequence h_n, we also have

$$\frac{\alpha_{h_n}^2}{1 - \alpha_{h_n} - \beta_{h_n}} \to \lambda \in [0, \infty].$$

The proof of the proposition is complete if the relation for β_h and the restrictions $\lambda \in (0,1)$ and $\rho = \lambda/(1 - \lambda)$ are proven. Similar to the calculations above we obtain

$$a(\alpha_{h_n}, \beta_{h_n}, \kappa_{h_n}, h/h_n) \to h\theta\lambda + h^2\theta^2\lambda/\rho + 2\lambda\{\exp(-h\theta) - 1 + h\theta\}, \qquad (B.3)$$

$$b(\alpha_{h_n}, \beta_{h_n}, h/h_n) \to h\theta\lambda\frac{1 - \exp(-2h\theta)}{2h\theta}, \qquad (B.4)$$

where the functions a and b are given in Theorem A.1. Suppose that $\lambda = 0$ (thus $\rho = 0$). Then the limit of the b function equals zero while the limit of the a function is still positive (possibly infinite). Hence the aggregation formula for β in Theorem A.1 implies $\beta_h = \lim_{n \to \infty} \alpha_{h_n} + \beta_{h_n} = 1$ for all $h > 0$. This is in contradiction with the weak GARCH assumption. Using $\lambda > 0$ and the aggregation formula for κ, this implies that $\kappa_{h_n} \to \kappa_0 \in [0, \infty)$ as $n \to \infty$, and, moreover, that $\rho = (\kappa_0 - 1)\lambda/(2 + \lambda)$. Plugging in this value for ρ into (B.2) and taking the limit on both sides for $h \downarrow 0$ yields $\kappa_0 = 3/(1 - \lambda)$ and hence $\lambda < 1$. This yields the desired values for ρ and κ_h. Finally insert ρ into the limit of the a function in (B.3) and obtain the value c_h. This proves the parameter configuration for GARCH diffusions. $\square$

Remark B.1. Suppose $\{Y_t, t \geqslant 0\}$ has continuous sample paths. Then a continuity condition on $f(h) = E|Y_h - Y_0|^4$ is equivalent to Assumption A. Proof: let h_n be a sequence with $h_n \to h$ as $n \to \infty$. Since $f(h_n) \to f(h)$ we obtain from Theorem 5.4 in Billingsley (1968) that the sequence $|Y_{h_n} - Y_0|^4$ is uniformly integrable. Using

$$|ab|I_{\{|ab| > L\}} \leqslant |a|^2 I_{\{|a| > \sqrt{L}\}} + |b|^2 I_{\{|b| > \sqrt{L}\}},$$

this also implies uniform integrability of $|(Y_{h_n} - Y_0)(Y_{(m+1)h_n} - Y_{mh_n})|^2$ for each $m \in \mathbb{N}$. Since $Y_h - Y_0$ is continuous the autocovariances of $\{|Y_{t+h} - Y_t|^2, t \in h\mathbb{N}\}$ are continuous functions of h. Finally, as the parameter $\zeta_h = (\psi_h, \alpha_h, \beta_h, \kappa_h)$ is a continuous function of the autocovariances, we obtain the desired continuity of ζ_h.

Proof of Proposition 4.1. This proof is completely similar to the proof of Proposition 3.1 by requiring

$$\frac{\alpha_{h_n}}{1 - \alpha_{h_n} - \beta_{h_n}} \to \phi \in [0, \infty].$$

(Re-)consideration of the sequences $h_n^{-1}a(\alpha_{h_n}, \beta_{h_n}, \kappa_{h_n}, h/h_n)$ and $h_n^{-1}b(\alpha_{h_n}, \beta_{h_n}, h/h_n)$ proves the result. The details are omitted. $\square$

Appendix C: Proofs of examples

Proof of Examples 3.1 and 4.1. We consider solutions of

$$dY_t = \sigma_{t-}\, dL_t, \tag{C.1}$$

$$d\sigma_t^2 = \theta(\omega - \sigma_{t-}^2)\, dt + \sqrt{2\eta\theta}\,\sigma_{t-}^2\, dM_t, \tag{C.2}$$

where $\{L_t, t \geq 0\}$ and $\{M_t, t \geq 0\}$ are independent standardized Lévy processes such that $\{L_t\}$ is symmetric, $EL_t^2 = EM_t^2 = t$, and $\kappa_t^L = EL_t^4/t^2$ exists and where $\omega > 0$, $\theta > 0$, and $\eta \in (0,1)$. We consider covariance stationary solutions of $\{\sigma_t^2, t \geq 0\}$, i.e., $E\sigma_t^2$ and $E\sigma_t^4$ are constant over time. Let $Y_0 = 0$ and define the filtration $\mathscr{F}_t = \mathscr{F}(Y_0, \sigma_0^2, L_s, M_s,\ s \in (0, t])$, $t \geq 0$, and the σ-field $\mathscr{G} = \mathscr{F}(\sigma_0^2, M_s,\ s > 0)$.

To show that the solution $\{Y_t, t \geq 0\}$ is GARCH according to Definition 2.2 we need to show that all discrete difference processes are weak GARCH or, equivalently, that the squares of the differences follow an ARMA(1,1) process. To prove the latter statement we will show that, for each $m \in \mathbb{N}$ and $h \in \mathbb{R}$,

$$\operatorname{cov}\{(Y_t - Y_{t-h})^2,\ (Y_{t-mh} - Y_{t-(m+1)h})^2\} = C_h(\omega, \theta, \eta) \exp(-mh\theta), \tag{C.3}$$

the autocovariances are exponentially decaying.

First we derive some results for $\{\sigma_t^2\}$. Using Fatou, the martingale property of $\{M_t\}$, the definition of $[\cdot,\cdot]$ on p. 58 of Protter (1990) and Theorem II.20, ibid., and Exercise 1.5.20 of Karatzas and Shreve (1988), we obtain, for fixed $t \geq 0$,

$$E(M_t - M_{t-})^2 \leq \liminf_{h \downarrow 0} E(M_t - M_{t-h})^2 = \liminf_{h \downarrow 0} E(M_t^2 - M_{t-h}^2)$$

$$= \liminf_{h \downarrow 0} E([M,M]_t - [M,M]_{t-h}) = 0.$$

Hence, by Theorem II.13 of Protter (1990), $\sigma_t^2 = \sigma_{t-}^2$ (a.s.) and, thus, for $s \leq t$,

$$E\left(\sigma_t^2 \,|\, \mathscr{F}_s\right) = E\left(\sigma_{t-}^2 \,|\, \mathscr{F}_s\right),$$

$$E\left(\sigma_t^4 \,|\, \mathscr{F}_s\right) = E\left(\sigma_{t-}^4 \,|\, \mathscr{F}_s\right).$$

Furthermore using Fubini's theorem we obtain, for $s \leq t$,

$$E\left(\sigma_t^2 \,|\, \mathscr{F}_s\right) - \sigma_s^2 = E\left(\int_{(s,t]} d\sigma_u^2 \,|\, \mathscr{F}_s\right)$$

$$= E\left(\int_{(s,t]} \theta(\omega - \sigma_{u-}^2)\, du \,|\, \mathscr{F}_s\right)$$

$$= \int_{(s,t]} \theta\left(\omega - E\left(\sigma_{u-}^2 \,|\, \mathscr{F}_s\right)\right) du.$$

Solving this differential equation yields, for $s \leqslant t$,

$$E\left(\sigma_t^2 \,|\, \mathscr{F}_s\right) = E\left(\sigma_{t-}^2 \,|\, \mathscr{F}_s\right) = \omega + (\sigma_s^2 - \omega)\exp\{-(t-s)\theta\}, \tag{C.4}$$

implying $E\sigma_t^2 = E\sigma_{t-}^2 = \omega$. Using Theorem II.29 of Protter (1990) and (C.4) we obtain

$$E\sigma_0^4 = E\sigma_t^4 = E\sigma_{t-}^4$$

$$= E\sigma_0^4 + E[\sigma^2 - \sigma_0^2, \sigma^2 - \sigma_0^2]_t + 2E\int_{(0,t]} \sigma_{u-}^2 \, d\sigma_u^2$$

$$= E\sigma_0^4 + 2\eta\theta \int_{(0,t]} E\sigma_{u-}^4 \, du + 2\int_{(0,t]} E\{\sigma_{u-}^2 \theta(\omega - \sigma_{u-}^2)\} du$$

$$= E\sigma_0^4 + 2t\theta\omega^2 + 2t\theta(\eta - 1)E\sigma_0^4$$

$$= \omega^2/(1 - \eta),$$

$$E\sigma_t^2\sigma_s^2 = \omega^2 + \omega^2 \frac{\eta}{1 - \eta} \exp(-|t - s|\theta). \tag{C.5}$$

These relations about $\{\sigma_t^2\}$ can be used to establish, using Theorem II.29 of Protter (1990) once more,

$$E(Y_t - Y_{t-h})^2 = \int_{(t-h,t]} E\sigma_{u-}^2 \, du = h\omega$$

and, by repeated use of the arguments above,

$$E\{(Y_t - Y_{t-h})^2(Y_{t-mh} - Y_{t-(m+1)h})^2\}$$

$$= E\{E\left((Y_t - Y_{t-h})^2 \,|\, \mathscr{F}_{t-mh}\right)(Y_{t-mh} - Y_{t-(m+1)h})^2\}$$

$$= E\left\{\int_{(t-h,t]} \sigma_{u-}^2 \, du E\left(Y_{t-mh} - Y_{t-(m+1)h}\right)^2 \,|\, \mathscr{G}\right)\right\}$$

$$= E\left\{E\left(\int_{(t-h,t]} \sigma_{u-}^2 \, du \,|\, \mathscr{F}_{t-mh}\right) \int_{(t-(m+1)h, t-mh]} \sigma_{u-}^2 \, du\right\}$$

$$= E\left\{(h\omega + (\sigma_{t-mh}^2 - \omega)\frac{\exp(h\theta) - 1}{\theta}\exp(-mh\theta)) \int_{(t-(m+1)h, t-mh]} \sigma_{u-}^2 \, du\right\}$$

$$= h^2\omega^2 + \frac{\exp(h\theta) - 1}{\theta}\exp(-mh\theta) \int_{(t-(m+1)h, t-mh]} E\{(\sigma_{t-mh}^2 - \omega)\sigma_{u-}^2\} du$$

$$= h^2\omega^2 + \frac{\exp(h\theta) - 1}{\theta}\exp(-mh\theta)\omega^2 \frac{\eta}{1 - \eta}\frac{1 - \exp(-h\theta)}{\theta}.$$

Hence (C.3) follows. The process defined by (C.1) and (C.2) is GARCH with parameter, say, $\zeta_h^Y = (\psi_h^Y, \alpha_h^Y, \beta_h^Y, \kappa_h^Y)$. In particular this implies that $\{L_t\}$ is GARCH with parameter $\zeta_h^L = (\psi_h^L, \alpha_h^L, \beta_h^L, \kappa_h^L) = (h, 0, 0, 3 + v_L^*/h)$.

Next we will derive the characterizing coefficients of $\{Y_t\}$. Two values are easily derived from the relations above:

$$\frac{\psi_h^Y}{1 - \alpha_h^Y - \beta_h^Y} = E(Y_t - Y_{t-h})^2 = h\omega,$$

$$\alpha_h^Y + \beta_h^Y = \frac{\text{cov}\{(Y_t - Y_{t-h})^2, (Y_{t-2h} - Y_{t-3h})^2\}}{\text{cov}\{(Y_t - Y_{t-h})^2, (Y_{t-h} - Y_{t-2h})^2\}} = \exp(-h\theta).$$

It remains to derive λ if $\{L_t\}$ is Brownian motion and v and ϕ in all other situations. To obtain these parameters we will consider the kurtosis of $Y_t - Y_{t-h}$. Define

$$S_n = \sum_{i=0}^{2^n-1} \sigma_{t-h+ih2^{-n}} (L_{t-h+(i+1)h2^{-n}} - L_{t-h+ih2^{-n}}), \qquad n \in \mathbb{N}.$$

If $\{S_n\}$ is a Cauchy sequence in $\mathscr{L}_4$, then, by Theorem II.21 of Protter (1990), $S_n \xrightarrow{4} Y_t - Y_{t-h}$. Hence, using (C.5),

$$\kappa_h^Y = E(Y_t - Y_{t-h})^4 / h^2 \omega^2$$

$$= \lim_{n \to \infty} E S_n^4 / h^2 \omega^2$$

$$= \lim_{n \to \infty} \sum_{i=0}^{2^n-1} E\{\sigma_{t-h+ih2^{-n}}^4 (L_{t-h+(i+1)h2^{-n}} - L_{t-h+ih2^{-n}})^4\} / h^2 \omega^2$$

$$+ 6 \lim_{n \to \infty} \sum_{i=0}^{2^n-1} \sum_{j=i+1}^{2^n-1} E\{\sigma_{t-h+ih2^{-n}}^2 \sigma_{t-h+jh2^{-n}}^2 (L_{t-h+(i+1)h2^{-n}} - L_{t-h+ih2^{-n}})^2$$

$$\times (L_{t-h+(j+1)h2^{-n}} - L_{t-h+jh2^{-n}})^2\} / h^2 \omega^2$$

$$= \lim_{n \to \infty} \sum_{i=0}^{2^n-1} (1 - \eta)^{-1} \kappa_{h2^{-n}}^L 2^{-2n}$$

$$+ 6 \lim_{n \to \infty} \sum_{i=0}^{2^n-1} \sum_{j=i+1}^{2^n-1} \left\{ 1 + \frac{\eta}{1 - \eta} \exp(-(j - i)h2^{-n}\theta) \right\} 2^{-2n}$$

$$= 3 + \frac{v_L^*}{(1 - \eta)h} + 6 \frac{\eta}{1 - \eta} \frac{\exp(-h\theta) - 1 + h\theta}{h^2\theta^2},$$

and we obtain the desired relationships for λ, v, and ϕ.

To complete the proof we establish $E|S_n - S_m|^4 \to 0$. Note, for $m > n$,

$$S_n - S_m = \sum_{i=0}^{2^n-1} \sum_{j=0}^{2^{m-n}-1} (\sigma_{t-h+ih2^{-n}} - \sigma_{t-h+ih2^{-n}+jh2^{-m}})$$

$$\times (L_{t-h+ih2^{-n}+(j+1)h2^{-m}} - L_{t-h+ih2^{-n}+jh2^{-m}}).$$

Using, for $a, b \geqslant 0$, $(a-b)^4 \leqslant (a^2 - b^2)^2$, (C.5), and Cauchy–Schwarz we obtain

$$
E|S_n - S_m|^4 = \sum_{i=0}^{2^n-1} \sum_{j=0}^{2^{m-n}-1} E(\sigma_{t-h+ih2^{-n}} - \sigma_{t-h+ih2^{-n}+jh2^{-m}})^4 \kappa_{h2^{-m}}^L h^2 2^{-2m}
$$

$$
+6 \sum_{i=0}^{2^n-1} \sum_{j=0}^{2^{m-n}-1} \sum_{p=i}^{2^n-1} \sum_{q=(j+1)\delta_{i..p}}^{2^{m-n}-1} E\{(\sigma_{t-h+ih2^{-n}} - \sigma_{t-h+ih2^{-n}+jh2^{-m}})^2
$$

$$
\times (\sigma_{t-h+ph2^{-n}} - \sigma_{t-h+ph2^{-n}+qh2^{-m}})^2\} h^2 2^{-2m}
$$

$$
\leqslant h^2 \omega^2 \frac{\eta}{1-\eta} 2^{n-m} \sum_{j=0}^{2^{m-n}-1} (1 - \exp(-jh2^{-m})) \left\{ 3 + \frac{v_L^*}{h2^{-m}} \right\} 2^{-m}
$$

$$
+ 3h^2 \omega^2 \frac{\eta}{1-\eta} \left\{ 2^{n-m} \sum_{j=0}^{2^{m-n}-1} (1 - \exp(-jh2^{-m}))^{1/2} \right\}^2
$$

$$
\to 0,
$$

as $m \geqslant n \to \infty$. $\square$

Appendix D: Some additional formulas

The coefficients a_h and b_h in Eq. (5.4) for GARCH diffusions and GARCH jump-diffusions are determined by

$$
a_h = -a_h^c / d_h^c, \qquad b_h = b_h^c / d_h^c,
$$

and

$$
a_h = -a_h^j / d_h^j, \qquad b_h = b_h^j / d_h^j,
$$

respectively, where

$$
a_h^c = 2h\lambda \frac{\exp\{-h(\theta + 2\tau)\} - \exp(-4h\tau) + h(\theta - 2\tau)\frac{1 - \exp(-4h\tau)}{4h\tau}}{h^2(\theta + 2\tau)(\theta - 2\tau)}
$$

$$
+ h(1 - \lambda) \left(\frac{1 - \exp(-2h\tau)}{2h\tau} \right)^2,
$$

$$
a_h^j = 2h\theta\phi(2 + \phi) \frac{\exp\{-h(\theta + 2\tau)\} - \exp(-4h\tau) + h(\theta - 2\tau)\frac{1 - \exp(-4h\tau)}{4h\tau}}{h^2(\theta + 2\tau)(\theta - 2\tau)}
$$

$$
+ 2h\theta v^{-1} \left(\frac{1 - \exp(-2h\tau)}{2h\tau} \right)^2 + \frac{1 - \exp(-4h\tau)}{4h\tau},
$$

$$b_h^c = \frac{\lambda\theta}{(\theta - 2\tau)^2} \left\{ \{1 + \exp(-2h\theta)\} \frac{1 - \exp(-4h\tau)}{4h\tau} \right.$$

$$+ \{1 + \exp(-4h\tau)\} \frac{1 - \exp(-2h\theta)}{2h\theta}$$

$$\left. - 2[1 + \exp\{-h(\theta + 2\tau)\}] \frac{1 - \exp\{-h(\theta + 2\tau)\}}{h(\theta + 2\tau)} \right\},$$

$$b_h^j = \frac{\theta^2\phi^2}{(\theta - 2\tau)^2} \left\{ \{1 + \exp(-2h\theta)\} \frac{1 - \exp(-4h\tau)}{4h\tau} \right.$$

$$+ \{1 + \exp(-4h\tau)\} \frac{1 - \exp(-2h\theta)}{2h\theta}$$

$$\left. - 2[1 + \exp\{-h(\theta + 2\tau)\}] \frac{1 - \exp\{-h(\theta + 2\tau)\}}{h(\theta + 2\tau)} \right\}$$

$$+ \frac{\theta\phi}{(\theta - 2\tau)} \left\{ 2\{1 + \exp(-2h\theta)\} \frac{1 - \exp(-4h\tau)}{4h\tau} \right.$$

$$\left. - 2[1 + \exp\{-h(\theta + 2\tau)\}] \frac{1 - \exp\{-h(\theta + 2\tau)\}}{h(\theta + 2\tau)} \right\},$$

$$d_h^c = \frac{\lambda\theta}{(\theta - 2\tau)^2} \left\{ \exp(-h\theta) \left(\frac{1 - \exp(-4h\tau)}{4h\tau} - \frac{1 - \exp\{-h(\theta + 2\tau)\}}{h(\theta + 2\tau)} \right) \right.$$

$$\left. + \exp(-2h\tau) \left(\frac{1 - \exp(-2h\theta)}{2h\theta} - \frac{1 - \exp\{-h(\theta + 2\tau)\}}{h(\theta + 2\tau)} \right) \right\},$$

$$d_h^j = \frac{\theta^2\phi^2}{(\theta - 2\tau)^2} \left\{ \exp(-h\theta) \left(\frac{1 - \exp(-4h\tau)}{4h\tau} - \frac{1 - \exp\{-h(\theta + 2\tau)\}}{h(\theta + 2\tau)} \right) \right.$$

$$\left. + \exp(-2h\tau) \left(\frac{1 - \exp(-2h\theta)}{2h\theta} - \frac{1 - \exp\{-h(\theta + 2\tau)\}}{h(\theta + 2\tau)} \right) \right\}$$

$$+ \frac{\theta\phi}{(\theta - 2\tau)} \left\{ 2\exp(-h\theta) \frac{1 - \exp(-4h\tau)}{4h\tau} \right.$$

$$\left. - \{\exp(-h\theta) + \exp(-2h\tau)\} \frac{1 - \exp\{-h(\theta + 2\tau)\}}{h(\theta + 2\tau)} \right\}.$$

If zero's appear in the denominators due to $\tau = 0$ or $\theta = 2\tau$, then one should take the corresponding limits to obtain a_h and b_h.

References

Amin, K.I., 1993, Jump diffusion option valuation in discrete time, Journal of Finance 48, 1833–1863.
Amin, K.I. and V.K. Ng, 1993, Option valuation with systematic stochastic volatility, Journal of Finance 48, 881–910.

Baillie, R.T. and T. Bollerslev, 1989, The message in daily exchange rates: A conditional-variance tale, Journal of Business and Economic Statistics 7, 297–305.

Billingsley, P. 1968, Convergence of probability measures (Wiley, New York, NY).

Black, F. and M. Scholes, 1973, The pricing of options and corporate liabilities, Journal of Political Economy 3, 637–654.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31, 309–328.

Bollerslev, T., R.Y. Chou, and K.F. Kroner, 1992, ARCH modeling in finance: A review of the theory and empirical evidence, Journal of Econometrics 52, 5–59.

Cox, J., J. Ingersoll, and S. Ross, 1985a, An intertemporal general equilibrium model of asset prices, Econometrica 53, 363–384.

Cox, J., J. Ingersoll, and S. Ross, 1985b, A theory of the term structure of interest rates, Econometrica 53, 385–407.

Diebold, F.X., 1988, Empirical modeling of exchange rates (Springer Verlag, New York, NY).

Drost, F.C. and C.A.J. Klaassen, 1996, Efficient estimation in semiparametric GARCH models, Journal of Econometrics, forthcoming.

Drost, F.C. and T.E. Nijman, 1993, Temporal aggregation of GARCH processes, Econometrica 61, 909–927.

Drost, F.C., T.E. Nijman, and B.J.M. Werker, 1994, Estimation and testing in models containing both jumps and conditional heteroskedasticity, CentER discussion paper 94105 (Tilburg University, Tilburg)..

Duffie, D. and K.J. Singleton, 1993, Simulated moments estimation of diffusion models of asset prices, Econometrica 61, 929–952.

Engle, R.F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, Econometrica 50, 987–1008.

Gallant, A.R. and G. Tauchen, 1992, Which moments to match?, Working paper (Duke University, Durham, NC).

Gourieroux, C., A. Monfort, and E. Renault, 1992, Indirect inference, Journal of Applied Econometrics 8, 85–108.

Heston, S.L., 1993, Invisible parameters in option prices, Journal of Finance 48, 933–947.

Huang, C., 1987, An intertemporal general equilibrium asset pricing model: The case of diffusion information, Econometrica 55, 117–142.

Johnson, H. and D. Shanno, 1987, Option pricing when the variance is changing, Journal of Financial and Quantitative Analysis 22, 143–151.

Jorion, P., 1988, On jump processes in the foreign exchange and stock markets, Review of Financial Studies 1, 427–445.

Karatzas, I. and S.E. Shreve, 1988, Brownian motion and stochastic calculus (Springer-Verlag, New York, NY).

Linton, O., 1993, Adaptive estimation in ARCH models, Econometric Theory 9, 539–569.

Melino, A. and S.M. Turnbull, 1990, Pricing foreign currency options with stochastic volatility, Journal of Econometrics 45, 239–265.

Merton, R.C., 1990, Continuous-time finance (Blackwell, Cambridge).

Naik, V., 1993, Option valuation and hedging strategies with jumps in the volatility of asset returns, Journal of Finance 48, 1969–1984.

Nelson, D.B., 1990, ARCH models as diffusion approximations, Journal of Econometrics 45, 7–38.

Nelson, D.B. and C.Q. Cao, 1992, Inequality constraints in the univariate GARCH model, Journal of Business and Economic Statistics 10, 229–235.

Nelson, D.B. and D.P. Foster, 1994, Asymptotic filtering theory for univariate ARCH models, Econometrica 62, 1–41.

Nijman, T.E. and E. Sentana, 1993, Marginalization and contemporaneous aggregation in multivariate GARCH processes, CentER discussion paper 9312 (Tilburg University, Tilburg); Journal of Econometrics, forthcoming.

Protter, P., 1990, Stochastic integration and differential equations (Springer-Verlag, New York, NY).

Revuz, D. and M. Yor, 1991, Continuous martingales and Brownian motion (Springer-Verlag, Berlin).

Scott, L., 1987, Option pricing when the variance changes randomly: Theory, estimation and an application, Journal of Financial and Quantitative Analysis 22, 419–438.

Vlaar, P.J.G. and F.C. Palm, 1993, The message in weekly exchange rates in the European monetary system: Mean reversion, conditional heteroscedasticity, and jumps, Journal of Business and Economic Statistics 11, 351–360.

Weiss, A.A., 1986, Asymptotic theory for ARCH models: estimation and testing, Econometric Theory 2, 107–131.