The COGARCH: A Review, with News on Option Pricing and Statistical Inference

Claudia Klüppelberg, Ross Maller and Alexander Szimayer

Abstract. Continuous time models have been elevated to great importance in the modelling of time series data, in response to the successful options pricing model of Black and Scholes (1973), among other things. In 2004, Klüppelberg, Lindner, and Maller introduced the "COGARCH" model as a continuous-time analogue to the enormously influential and successful discrete time GARCH stochastic volatility model of Engle and Bollerslev. Like the GARCH model, the COGARCH is based on a single source of random variability, in this case, on a single background driving Lévy process.

Since its inception, the original COGARCH model has been studied intensively and generalised and extended in various ways. In the present paper we formulate the model using stochastic differential equations and review some of its important properties as well as some recent developments, including some statistical issues. As a new contribution we present a COGARCH option pricing model including the possibility of default, in which the underlying stock price process is taken as a stochastic exponential of a COGARCH model with drift. We give a preliminary analysis of this model in its risk-neutral dynamics, and as a prominent example, compute European option prices in the Variance-Gamma COGARCH model.

For practical implementation, we must discretise the continuous-time COGARCH onto a discrete grid over a finite time interval. We go on to review ways of doing this by means of various approximation schemes, in particular, via a "first jump" approximation to the underlying Lévy process, which preserves features of the process important for optimal stopping problems. Some other applications of the technology, especially, to the modelling of irregularly spaced time series data, are discussed too.

2010 Mathematics Subject Classification. Primary 60G51, 60J25, 62M05, 91B25, 91B84; Secondary 37A50, 60F17, 60G10, 91B70.

Keywords. Continuous-time GARCH model, Lévy process, option pricing, stochastic volatility, statistics for stochastic processes, functional limit theorems.

1. Introduction

Mathematical Finance and Econometrics can be viewed as two sides of a coin. Econometrics concentrates on finding optimal models concerning statistical properties like correlations and prediction. Mathematical Finance, on the other side is mainly concerned with finding good models which allow for hedging and derivatives pricing.

Two major innovations revolutionised the theory and practise of econometrics in the latter part of the last century. The first was the development of the unit root and related cointegration concepts in the analysis of time series data, and the associated Dickie-Fuller test (cf. [11]) and its various generalisations. Soon after came the idea of conditional heteroscedasticity models to capture the empirically observed feature of apparently randomly varying volatility[1] fluctuations in time series. Landmark models include Taylor's stochastic volatility model [47] and the ARCH and GARCH models of Engle [16] and Bollerslev [4]. These innovations and their subsequent rapid development and application in many directions were particularly appropriate for high frequency time series financial data, which became easily accessible in large quantities over the period, with the introduction of modern computer technology.

Rigorous hedging and pricing of financial derivatives started with the seminal paper by Black and Scholes [3] using the complete model of geometric Brownian motion and its explicit unique option price. After it became clear that this model cannot capture all realistic features of market conditions, incomplete models entered the scene. Characterization of no arbitrage pricing by martingale measures came into focus in the important papers by Harrison and Kreps [22] and Harrison and Pliska [23]. The problem of non-unique martingale measures was met by a specific approach of Föllmer and Schweizer [18]. Exponential Lévy models were a first step towards more realistic modelling promoted early on by Eberlein and collaborators; cf. Eberlein [15] for a review. Pricing measures were suggested for normal mixture models such as the variance gamma model due to Madan and Seneta [35] and the normal inverse Gaussian model, which was originally suggested by Barndorff-Nielsen [1].

This paper aims at a reconciliation between certain econometric models and pricing models. Our econometrics motivation comes from the availability of high frequency data, which are often sampled at irregular time points, making continuous-time modelling necessary. Our motivation for derivatives pricing originates in the need for more realistic pricing models.

The paper is organised as follows. In Section 2 we recall the discrete time ARCH and GARCH models and some of their properties. We further summarize some continuous-time limits of such models from the literature and explain their drawbacks. Section 3 is devoted to the continuous time GARCH (COGARCH) model as suggested in Klüppelberg, Lindner and Maller [30]. Section 4 presents new material on option pricing within the COGARCH model. As an explicit example we treat the variance gamma driven COGARCH and compare it to the Heston model via implicit volatilities. It turns out that the COGARCH can produce higher implied volatilities for short maturities deep in-the-money and far out of-the-money; desirable properties in applications. Section 5 is devoted to statistical estimation of the COGARCH parameters. Besides classical moment estimators we also present a method to obtain a GARCH skeleton within the COGARCH model, which allows for the use of existing software for extimation. This involves functional limit theorems in various modes of convergence.

[1] The "volatility" is simply the square root of the variance.

2. Background in Discrete Time ARCH and GARCH Models

AutoRegressive Conditional Heteroscedasticity (ARCH) models were introduced by Engle [16] and soon generalised to GARCH (Generalised AutoRegressive Conditional Heteroscedasticity) models by Bollerslev [4]. Nowadays they are seen as particular kinds of stochastic volatility models, in which the variance of time series innovations is itself assumed to vary randomly, conditional on past information.

A special feature of ARCH and GARCH models is that they incorporate feedback between an observation and its volatility, whereby a large fluctuation in an innovation triggers a corresponding large fluctuation in the variance of the series, which in the absence of further large fluctuations, then reverts to a steady state level, as long as the process is in a stationary regime. This is an attractive concept, which accords well with intuition and empirical observation of, especially, financial time series. As it turns out, the models also display further desirable features from the modelling point of view. In particular, they typically induce long tailedness of marginal distributions, and serial correlations, not in the innovations themselves, but in the *squared* innovations. These features again accord well with empirical observation. We expand further on them later.

The simplest GARCH model, the GARCH(1,1), is a discrete time process with three parameters, $\beta > 0$, $\phi \geq 0$, $\delta \geq 0$, specifying the variance as a discrete time stochastic recursion, or difference equation. We write it using two equations, one specifying the "mean level" process (the observed data, perhaps after removal of trend or other deterministic feature, to approximate stationarity) and the other specifying the variance process, which is time dependent and randomly fluctuating. Thus, for $i = 1, 2, \ldots$,

$$Y_i = \varepsilon_i \sigma_i, \tag{2.1}$$

with

$$\sigma_i^2 = \beta + \phi Y_{i-1}^2 + \delta \sigma_{i-1}^2 = \beta + (\phi \varepsilon_{i-1}^2 + \delta)\sigma_{i-1}^2. \tag{2.2}$$

Here the starting values ε_0 and σ_0 are given quantities, possibly random, and usually assumed independent of the $(\varepsilon_i)_{i=1,2,\ldots}$, which are the sole source of variation in the model. The ε_i, $i = 1, 2, \ldots$ are assumed to be independent identically distributed (i.i.d.) random variables (rvs) centered at 0. Serial dependence between the Y_i is introduced via the dependence of the σ_i^2 on their past values. Conditional on σ_i, Y_i simply has the distribution of ε_i, scaled by σ_i, which in general (as long as $\phi, \delta > 0$) is time dependent, hence the "conditional heteroscedasticity" part of the terminology. The "autoregressive" aspect refers to the form of the dependence of σ_i^2 on σ_{i-1}^2, as in an autoregressive time series model.

When $\delta = 0$ this term in σ_{i-1}^2 disappears, but volatility remains stochastic via the dependence of σ_i^2 on Y_{i-1}, as long as $\phi > 0$. We then have the ARCH(1) model of Engle [16]. Such a model was generally found to be inadequate, however, to describe observed data, in which variance tends to be highly persistent and mean reverting. The introduction of the σ_{i-1}^2 term in (2.2) when $\delta > 0$ improves the modelling of such data substantially, and gives rise to the highly successful GARCH(1,1) model of Bollerslev [4]. A very natural extension of this model is to add further autoregressive terms to (2.2), thus defining a GARCH(p, q) model, and, similarly, the ARCH(p) model is defined.

Of course if $\phi = \delta = 0$ in (2.2) the model simply reduces to one of i.i.d. observations for the Y_i, with variance $\beta > 0$.

2.1. Stationarity and Tail Behaviour in GARCH Models. Often, in a practical regression situation, the ε_i might be assumed $N(0, 1)$ for the purposes of model fitting. Such a short tailed distribution for the ε_i, however, does not necessarily translate into a short tailed marginal distribution for the observations Y_i. Eq. (2.2) specifies the sequence $(\sigma_i^2)_{i=1,2,\ldots}$ as a *stochastic recurrence equation*, studied in some depth in the probability literature, especially, see Kesten [28], Vervaat [49] and Goldie [19]; see also the readable overview paper by Diaconis and Freedman [10]. In a stationary regime, or otherwise, the resulting Y_i will usually have a heavy tailed distribution. This comes about as follows. Necessary and sufficient conditions for the "stability" (existence of an almost sure (a.s.) limit for large times) of a discrete time stochastic perpetuity given in Goldie and Maller [20] can be applied directly to give necessary and sufficient conditions in terms of log moments of the ε_i and the parameters ϕ and δ, for the stability of the ARCH(1) and GARCH(1,1) models. Specifically, Theorem 2.1 of Klüppelberg, Lindner and Maller [30] shows that we have stability of the mean and variance processes, that is, $Y_i \stackrel{D}{\to} Y_\infty$ and $\sigma_i \stackrel{D}{\to} \sigma_\infty$, as $i \to \infty$, for finite rvs Y_∞ and σ_∞, if and only if

$$E|\log(\delta + \phi\varepsilon_1^2)| < \infty \quad \text{and} \quad E\log(\delta + \phi\varepsilon_1^2) < 0.$$

These then constitute conditions for stationarity of $(Y_i, \sigma_i^2)_{i=1,2,\ldots}$ if the sequence is started with the values $(Y_\infty, \sigma_\infty)$. Then, further results transferable from the theory of stochastic difference equations show that, under certain fairly general conditions, Y_∞ will have a long tailed distribution, specifically, a distribution with a Pareto (power law) tail. A good exposition of this is in Lindner [32].

Thus, even with a short tailed distribution such as the normal assumed for the innovations ε_i, we may expect a heavy tailed marginal distribution for the Y_i. This accords with observed features of, especially, financial data, cf. Klüppelberg [29], Mikosch [37]. More recently, Platen and Sidorowicz [42], for example, in a very extensive investigation, suggest that much financial returns data has a very heavy tailed distribution, such as a t–distribution with 4 degrees of freedom.

2.2. Continuous Time Limits of GARCH Models. Motivated, in particular, by the availability of high-frequency data and by a need for option pricing technologies, classical diffusion limits have been used in a natural way to suggest continuous time limits of discrete time processes, including for the GARCH models. The best known of these is due to Nelson [40]. His limiting diffusion model is:

$$\mathrm{d}Y_t = \sigma_t \mathrm{d}B_t^{(1)},\ t \geq 0, \tag{2.3}$$

where σ_t, the volatility process, satisfies

$$\mathrm{d}\sigma_t^2 = \left(\beta - \eta\sigma_t^2\right)\mathrm{d}t + \phi\sigma_t^2 \mathrm{d}B_t^{(2)},\ t > 0, \tag{2.4}$$

with $B^{(1)}$ and $B^{(2)}$ independent Brownian motions, and $\beta > 0$, $\eta \geq 0$ and $\phi \geq 0$ constants.

Unfortunately, in these situations, the limiting models can lose certain essential properties of the discrete time GARCH models. It is surprising and counterintuitive, for example, that Nelson's diffusion limit of the GARCH process is driven by two independent Brownian motions, i.e. has two independent sources of randomness[2], whereas the discrete time GARCH process is driven only by a single white noise sequence. One of the features of the GARCH process is the idea that large innovations in the price process are almost immediately manifested as innovations in the volatility process; but this feedback mechanism is lost in models such as the Nelson continuous time version. Further, the appearance of an extra source of variation can have implications for completeness considerations in options pricing models, for example.

The phenomenon that a diffusion limit may be driven by two independent Brownian motions, while the discrete time model is given in terms of a single white noise sequence, is not restricted to the classical GARCH process. Duan [13] has shown that this occurs for many GARCH-like processes. On the other hand, Corradi [8] modified Nelson's method to obtain a diffusion limit depending only on a single Brownian motion - but then the equation for σ_t^2 degenerates to an ordinary differential equation. Using a Brownian bridge between discrete time observations, Kallsen and Taqqu [26] found a complete model driven by only one Brownian motion.

Moreover, the continuous time limits found in such a way can have distinctly different statistical properties to the original discrete time processes. As was shown by Wang [50], parameter estimation in the discrete time GARCH and the corresponding continuous time limit stochastic volatility model may yield different estimates (see also Brown, Wang and Zhao [6]). Thus these kinds of continuous time models are probabilistically and statistically different from their discrete time progenitors. See Lindner [31] for a recent overview of continuous time approximations to GARCH processes.

In Klüppelberg, Lindner and Maller [30], the authors proposed a radically different approach to obtaining a continuous time model. Their "COGARCH" (continuous time GARCH) model is a direct analogue of the discrete time GARCH, based on a single background driving Lévy process, and generalises the essential features of the discrete time GARCH process in a natural way. In the next section we review this model.

Generally, in what follows, by the "COGARCH" model we will mean the COGARCH(1,1) model.

[2]Dependence has been introduced in the literature in an ad hoc way by allowing $B^{(1)}$ and $B^{(2)}$ in (2.3) and (2.4) to be correlated, but such models still rely on two sources of randomness.

3. The COGARCH model

The COGARCH model is specified by two equations, the mean and variance equations, analogous to (2.1) and (2.2). The single source of variation is a so-called *background driving Lévy process* $L = (L_t)_{t\geq 0}$ with characteristic triplet (γ, σ^2, Π); we refer to Sato [43] for background on Lévy processes. The continuous time process L has i.i.d. increments, which are analogous to the i.i.d. innovations ε_i in (2.1) and (2.2). Then the COGARCH process $(G_t)_{t\geq 0}$ is defined in terms of its stochastic differential, $\mathrm{d}G$, such that

$$\mathrm{d}G_t = \sigma_{t-}\mathrm{d}L_t,\ t \geq 0, \tag{3.1}$$

where σ_t, the volatility process, satisfies

$$\mathrm{d}\sigma_t^2 = \left(\beta - \eta\sigma_{t-}^2\right)\mathrm{d}t + \phi\sigma_{t-}^2\mathrm{d}[L, L]_t,\ t > 0, \tag{3.2}$$

for constants $\beta > 0$, $\eta \geq 0$ and $\phi \geq 0$. Here $[L, L]_t$ denotes the quadratic variation process of L, defined for $t > 0$ by

$$[L, L]_t \quad := \quad \sigma^2 t + \sum_{0<s\leq t} (\Delta L_s)^2 \ = \ \sigma^2 t + [L_t, L_t]^d, \tag{3.3}$$

with $[L_t, L_t]^d$ denoting the pure jump component of $[L, L]$. (There should be no confusion between the constant σ^2 specifying the variance of the Gaussian component of L and the COGARCH variance process $(\sigma_t^2)_{t\geq 0}$. In (3.3), $\Delta L_t = L_t - L_{t-}$ for $t \geq 0$ (with $L_{0-} = 0$) and similarly for other processes throughout. All processes are cádlág)

To see the analogy with (2.1) and (2.2), note from (2.2) that

$$\sigma_i^2 - \sigma_{i-1}^2 = \beta - (1 - \delta)\sigma_{i-1}^2 + \phi\sigma_{i-1}^2\varepsilon_{i-1}^2, \tag{3.4}$$

which corresponds to (3.2) (with a reparameterisation from η to $\delta = 1 - \eta$) when the time increment $\mathrm{d}t$ is taken as a unit, or at least fixed, interval of time. But an advantage of the continuous time setup is that non-equally spaced observations are easily catered for, as we demonstrate later (Section 5.4).

Just as an understanding of discrete time perpetuities is the key to stability, stationarity and tail behaviour of the discrete time GARCH, so kinds of continuous time perpetuities are instrumental in the analysis of the COGARCH. The solution of the stochastic differential equation (SDE) (3.2) can be obtained with the help of an auxiliary Lévy process $X = (X_t)_{t\geq 0}$ defined by

$$X_t = \eta t - \sum_{0<s\leq t} \log(1 + \phi(\Delta L_s)^2),\ t \geq 0. \tag{3.5}$$

X is a spectrally negative Lévy process of bounded variation arising in a natural way in Klüppelberg et al. [30], where the COGARCH(1,1) is motivated directly as an analogue to the discrete time GARCH(1,1) process. Using Ito's lemma, it can be verified that the solution of (3.2) can be written in terms of X as

$$\sigma_t^2 = e^{-X_t}\left(\beta \int_0^t e^{X_s}\mathrm{d}s + \sigma_0^2\right),\ t \geq 0, \tag{3.6}$$

which reveals σ_t^2 as a kind of generalised Ornstein-Uhlenbeck (GOU) process, parameterised by (β, η, φ), and driven by the Lévy process L. For results on the GOU, and associated studies of *Lévy integrals*, see Lindner and Maller [33] and their references. An understanding of stability, stationarity and tail behaviour properties for the GOU is essential for such issues relating to G.

Klüppelberg et al. [30], Theorem 3.2, shows that the variance process $(\sigma_t^2)_{t\geq 0}$ for the COGARCH is a time homogeneous Markov process, and, further, that the bivariate process $(G_t, \sigma_t^2)_{t\geq 0}$ is Markovian. A finite random variable σ_∞^2 exists as the limit in distribution of σ_t^2 as $t \to \infty$ if and only if

$$\int_{\mathbb{R}} \log\Big(1 + \frac{\phi}{\delta} y^2\Big)\Pi(\mathrm{d}y) < -\log\delta. \tag{3.7}$$

σ_∞^2 has the same distribution as β times the stochastic integral $\int_0^\infty e^{-X_t}\mathrm{d}t$, which exists as a finite rv a.s. under (3.7) (see Theorem 3.1 of [30]). If this is the case and $(\sigma_t^2)_{t\geq 0}$ is started with value σ_∞^2, i.e., σ_0^2 is taken to have the distribution of σ_∞^2, independent of L, then $(\sigma_t^2)_{t\geq 0}$ is strictly stationary and $(G_t)_{t\geq 0}$ is a process with stationary increments (Theorem 3.2 and Corollary 3.1 of [30]).

Moments of the COGARCH process can be calculated using the Laplace transform of the auxiliary process X, which satisfies $\mathbb{E}e^{-\theta X_t} = e^{t\Psi(\theta)}$, with

$$\Psi(\theta) = -\eta\theta + \int_{\mathbb{R}} \big((1+\phi x^2)^\theta - 1\big)\Pi(dx),\ \theta \geq 0\,. \tag{3.8}$$

Returns over time intervals of fixed length $r > 0$ we denote by

$$G_t^{(r)} := G_t - G_{t-r} = \int_{(t-r,t]} \sigma_{s-}\, dL_s,\ t \geq r\,, \tag{3.9}$$

so that $(G_{ri}^{(r)})_{i\in\mathbb{N}}$ describes an equidistant sequence of non-overlapping returns. Calculating the corresponding quantity for the volatility yields

$$\begin{aligned} \sigma_{ri}^{2(r)} &:= \sigma_{ri}^2 - \sigma_{r(i-1)}^2 = \int_{(r(i-1),ri]} \big((\beta - \eta\sigma_s^2)\,\mathrm{d}s + \varphi\,\sigma_{s-}^2\,\mathrm{d}[L,L]_s\big) \\ &= \beta r - \eta \int_{(r(i-1),ri]} \sigma_s^2\, ds + \varphi \int_{(r(i-1),ri]} \sigma_{s-}^2\, \mathrm{d}[L,L]_s\,. \end{aligned} \tag{3.10}$$

Note that the stochastic process

$$\int_{(0,t]} \sigma_{s-}^2\, \mathrm{d}[L,L]_s = \sigma^2 \int_0^t \sigma_{s-}^2 ds + \sum_{0<s\leq t} \sigma_{s-}^2 (\Delta L_s)^2,\ t \geq 0,$$

is the quadratic variation $[G,G]_t$ of G, which satisfies

$$[G,G]_t = \int_0^t \sigma_{s-}^2\, d[L,L]_s,\ t \geq 0;$$

thus $\int_{(r(i-1),ri]} \sigma_{s-}^2\, \mathrm{d}[L,L]_s^d$ in (3.10) corresponds to the jump part of the quadratic variation of G accumulated during $(r(i-1), ri]$.

The following result (Proposition 2.1 of Haug et al. [24]) shows that the COGARCH has a similar moment structure as the GARCH model; in particular, increments are uncorrelated, but squared increments are positively correlated. We shall need these results in Section 5.1, when we present a method of moment estimation of the COGARCH parameters.

Proposition 3.1. *Suppose that* $(L_t)_{t\geq 0}$ *has finite variance and zero mean, and that the Gaussian component has variance* σ^2*. Suppose also that* $\Psi(1) < 0$ *for* Ψ *as given in* (3.8)*. Let* $(\sigma_t^2)_{t\geq 0}$ *be the stationary volatility process, so that* $(G_t)_{t\geq 0}$ *has stationary increments. Then* $\mathbb{E}(G_t^2) < \infty$ *for all* $t \geq 0$*, and for every* $t, h \geq r > 0$ *we have*

$$\mathbb{E}(G_t^{(r)}) \;=\; 0\,, \quad \mathbb{E}(G_t^{(r)})^2 \;=\; \frac{\beta r}{|\Psi(1)|}\mathbb{E}L_1^2\,, \quad \mathrm{Cov}(G_t^{(r)}, G_{t+h}^{(r)}) \;=\; 0.$$

If, further, $\mathbb{E}(L_1^4) < \infty$ *and* $\Psi(2) < 0$*, then* $\mathbb{E}(G_t^4) < \infty$ *for all* $t \geq 0$ *and, if, additionally, the Lévy measure* Π *of* L *is such that* $\int_{\mathbb{R}} x^3 \Pi(dx) = 0$*, then for every* $t, h \geq r > 0$*, we have*

$$\mathbb{E}(G_t^{(r)})^4 \;=\; 6\mathbb{E}(L_1^2)\frac{\beta^2}{\Psi(1)^2}\left(\frac{2\eta}{\varphi} + 2\sigma^2 - \mathbb{E}L_1^2\right)\left(\frac{2}{|\Psi(2)|} - \frac{1}{|\Psi(1)|}\right),$$
$$\left(r - \frac{1-e^{-r|\Psi(1)|}}{|\Psi(1)|}\right) + \frac{2\beta^2}{\varphi^2}\left(\frac{2}{|\Psi(2)|} - \frac{1}{|\Psi(1)|}\right) r + \frac{3\beta^2}{\Psi(1)^2}(\mathbb{E}L_1^2)^2 r^2$$

and

$$\begin{aligned}\mathrm{Cov}((G_t^{(r)})^2, (G_{t+h}^{(r)})^2) \;&=\; \frac{\mathbb{E}(L_1^2)\beta^2}{|\Psi(1)|^3}\left(\frac{2\eta}{\varphi} + 2\sigma^2 - \mathbb{E}L_1^2\right)\left(\frac{2}{|\Psi(2)|} - \frac{1}{|\Psi(1)|}\right)\\ &\quad \left(1 - e^{-r|\Psi(1)|}\right)\left(e^{r|\Psi(1)|} - 1\right) e^{-h|\Psi(1)|} \;>\; 0\,.\end{aligned}$$

Motivated by the generalization of the GARCH(1,1) to the GARCH(p, q) model, Brockwell, Chadraa, and Lindner [5] introduced a COGARCH(p, q) model. In it, the volatility follows a CARMA (continuous-time ARMA) process driven by a Lévy process (cf. Doob [12], Todorov and Tauchen [48]).

In Stelzer [45] multivariate COGARCH(1,1) processes are introduced, constituting a dynamical extension of normal mixture models and again incorporating such features as dependence of returns (but without autocorrelation), jumps, heavy tailed distributions, etc. Stelzer's definition agrees for $d = 1$ with the COGARCH(1,1) process. As in the univariate case, the model has only one source of randomness, a single multivariate Lévy process. The time-varying covariance matrix is modelled as a stochastic process in the class of positive semi-definite matrices. In [45] Stelzer analyses the probabilistic properties of the model and gives a sufficient condition for the existence of a stationary distribution for the stochastic covariance matrix process, and criteria ensuring the finiteness of moments.

4. A COGARCH Option Pricing Model

A potentially important application of the COGARCH model is to option pricing. Traditionally, and for mathematical tractability, option pricing models are based on continuous time models for an underlying stock price process. The discrete-time GARCH reproduces features commonly observed in financial data, especially relating to the so-called stylized facts (volatility clustering, mean reversion of volatility, negative skew, and heavy tails). Consequently, the COGARCH, as a continuous time limit of the discrete GARCH (see Section 5.3), can be expected to result in more accurate option valuation than standard models. In this section we propose an option pricing framework, where the stock price return is driven by COGARCH, thus allowing for stochastic volatility, and we also include the possibility of default in the model. Combining these features is not new; however, our framework is parsimonious in its parameterisation and as we will see can reproduce observed kinds of volatility smile and skew quite well. Further, the default probability in the model can be expressed as a function of the volatility.[3]

The financial market is defined on a filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\geq 0})$ satisfying the usual hypothesis, which is large enough to support a Lévy process $L = (L_t)_{t\geq 0}$ with characteristic function given for every $t \geq 0$ by $E[e^{izL_t}] = e^{t\psi(z)}$ for $z \in \mathbb{R}$ where

$$\psi(z) = i\gamma z - \frac{\sigma^2}{2} z + \int_{\mathbb{R}\setminus\{0\}} \Big(1 - e^{izx} - izh(x)\Big)\, \Pi(\mathrm{d}x).$$

As usual (γ, σ^2, Π) is the characteristic triplet, with related truncation function $h(x) = x\,\mathbf{1}_{\{|x|\leq 1\}}$. As a technical prerequisite we assume that the fourth moment of L exists, i.e. $\int x^4\, \Pi(\mathrm{d}x) < \infty$. The investment opportunities considered here are the risk-free money market account and the risky company stock. The *risk-free money market account* has the price process $B = (B_t)_{t\geq 0}$ with dynamics

$$\mathrm{d}B_t = rB_t\, \mathrm{d}t\,, \quad B_0 = 1\,, \tag{4.1}$$

where $r \in \mathbb{R}$ is the instantaneous risk-free rate; hence $B_t = e^{r\,t}$ for $t \geq 0$. The *stock price process* is denoted by $S = (S_t)_{t\geq 0}$ and bears two kinds of related risks. The stock price fluctuation is driven by a COGARCH process $G = (G_t)_{t\geq 0}$ with its accompanying volatility process $(\sigma_t)_{t\geq 0}$, and, further, the stock price is assumed to fall to zero at a default time τ, if default occurs, after which it stays at that level. Before default the stock price process satisfies

$$\mathrm{d}S_t = S_{t-}\, \mathrm{d}R_t\,, \quad S_0 > 0\,, \tag{4.2}$$

where $R = (R_t)_{t\geq 0}$, the cumulative return process, is driven by the COGARCH process G in the following sense:

$$\mathrm{d}R_t = [r + \lambda(\sigma_{t-})\, \sigma_{t-}]\, \mathrm{d}t + \sigma_{t-}\, \mathrm{d}L_t. \tag{4.3}$$

[3] As is also the case with the expected default frequency in Moody's KMV (Kealhofer, McQuown and Vasicek) EDF (Expected Default Frequency) proprietry credit measures model, the KMV EDF$^{\mathrm{TM}}$; see, e.g., http://www.moodyskmv.com/newsevents/files/EDF_Overview.pdf

Here the scaled innovation $\sigma_{t-}\,\mathrm{d}L_t$ is a COGARCH increment, $\mathrm{d}G_t$, $\lambda : [0,\infty) \to \mathbb{R}$ is the risk premium, and the volatility $(\sigma_t^2)_{t\geq 0}$ follows the dynamics in (3.2), namely:

$$\mathrm{d}\sigma_t^2 = \left(\beta - \eta\,\sigma_{t-}^2\right)\mathrm{d}t + \phi\,\sigma_{t-}^2\,\mathrm{d}[L,L]_t,\ t>0. \tag{4.4}$$

The default time τ is defined as the first time at which the cumulative return R exhibits a jump ΔR_t below $-100\% = -1$:

$$\tau = \inf\{t>0 : \Delta R_t \leq -1\} = \inf\{t>0 : \sigma_{t-}\,\Delta L_t \leq -1\}\,.$$

At default the stock price drops to zero and stays there, thus we can write

$$S_t = S_0\,\mathcal{E}(R)_t\,\mathbf{1}_{\{t<\tau\}}\,, \tag{4.5}$$

where $\mathcal{E}(X)$ denotes the stochastic exponential of X.

4.1. Relationship to Other Stochastic Volatility Models. To compare our model to other SV models, we reparameterise (4.3) and (4.4) as follows. Let us first assume that L in (4.3) is an error term satisfying $\mathbb{E}L_t = 0$ and $\mathbb{E}[L_t^2] = t$, or, equivalently,

$$\gamma + \int_{|x|>1} x\,\Pi(\mathrm{d}x) = 0, \quad \text{and} \quad \sigma^2 + \int_{\mathbb{R}} x^2\,\Pi(\mathrm{d}x) = 1. \tag{4.6}$$

This assumption is in fact no restriction, but ensures that the parameters can be identified. (Note that the function λ can be adjusted when centering L, and the scaling to unit variance of L affects only the variance parameters.)

The bracket process $[L,L]$ drives the volatility process σ. We center and scale $[L,L]$ to a martingale M with unit variance rate

$$M_t = \frac{[L,L]_t - \mathbb{E}[L,L]_t}{\sqrt{\mathbb{E}\left[([L,L]_1 - \mathbb{E}[L,L]_1)^2\right]}} = \frac{[L,L]_t^d - t\int_{\mathbb{R}} x^2\,\Pi(\mathrm{d}x)}{\sqrt{\int_{\mathbb{R}} x^4\,\Pi(\mathrm{d}x)}}\,, \quad t\geq 0. \tag{4.7}$$

Then we can write the variance equation (4.4) as

$$\mathrm{d}\sigma_t^2 = \kappa\,\left(\overline{\sigma}^2 - \sigma_{t-}^2\right)\mathrm{d}t + \nu\,\sigma_{t-}^2\,\mathrm{d}M_t,\ t>0, \tag{4.8}$$

where

$$\kappa = \eta - \phi\,, \quad \overline{\sigma}^2 = \frac{\beta}{\eta-\phi} \quad \text{and} \quad \nu = \phi\sqrt{\int_{\mathbb{R}} x^4\,\Pi(\mathrm{d}x)}\,. \tag{4.9}$$

The variance process is thus seen to be mean-reverting with mean level $\overline{\sigma}^2$, mean-reversion speed κ, and volatility $(\nu\sigma_t^2)_{t\geq 0}$, implying an average volatility of the variance process of $\nu\,\overline{\sigma}^2$. This enables us to benchmark our model to other SV models. We compare the COGARCH with the stochastic volatility model of Heston [25] (other related models include a Heston extension allowing for jumps of

Bates [2], the SABR model of Hagan et al. [21], etc.). The dynamics of the Heston model are

$$\begin{aligned} \mathrm{d}S_t^H &= \mu_t^H\, S_t^H \mathrm{d}t + \sigma_t^H\, S_t^H\, \mathrm{d}W_t^{H,1}\,, \\ \mathrm{d}(\sigma_t^H)^2 &= \kappa^H\, \left((\bar\sigma^H)^2 - (\sigma_t^H)^2\right) \mathrm{d}t + \nu^H\, \sqrt{(\sigma_t^H)^2}\, \mathrm{d}W_t^{H,2}\,, \end{aligned}$$

with expected return rate μ^H, mean-reversion speed κ^H, mean level $(\bar\sigma^H)^2$, volatility of volatility parameter ν^H, and leverage ρ^H. Here, the leverage ρ^H is in fact the correlation of the standard Brownian motions $W^{H,1}$ and $W^{H,2}$.

Our model features the so-called option pricing leverage effect that is also included in the Heston model. However, in our setup leverage is not a free parameter, but is determined by the skew and kurtosis of the jump measure of L. Formally, leverage is quantified by the instantaneous correlation between the increments of the price equation $\mathrm{d}R_t$ and the increments of the variance equation $\mathrm{d}\sigma_t^2$. By scaling this reduces to the correlation of L_t and M_t, and leverage is given by

$$\rho = \mathrm{cor}(L_t, M_t) = \frac{\mathbb{E}[L_t\, M_t]}{t} = \frac{1}{t}\, \mathbb{E}\left[[L, M]_t\right] = \frac{\int_{\mathbb{R}} x^3\, \Pi(\mathrm{d}x)}{\sqrt{\int_{\mathbb{R}} x^4\, \Pi(\mathrm{d}x)}}\,. \tag{4.10}$$

We see that ρ is restricted by more than just $|\rho| \le 1$; the Cauchy-Schwarz inequality implies

$$|\rho| \le \sqrt{\int_{\mathbb{R}} x^2\, \Pi(\mathrm{d}x)} = \sqrt{1-\sigma^2} \quad \text{(cf. (4.6))}.$$

The COGARCH and Heston models are compared in Table 1.

	Drift	Volatility	Noise	Leverage
COGARCH	$\kappa\left(\overline{\sigma}^2 - \sigma_t^2\right)$	$\nu\, \sigma_t^2$	f.v. pure jump	$\frac{\int_{\mathbb{R}} x^3\, \Pi(\mathrm{d}x)}{\sqrt{\int_{\mathbb{R}} x^4\, \Pi(\mathrm{d}x)}}$
Heston	$\kappa\left(\overline{\sigma}^2 - \sigma_t^2\right)$	$\nu\, \sqrt{\sigma_t^2}$	Brownian	$\rho \in [-1, 1]$

Table 1. Specifications of the variance processes for COGARCH and Heston models ("f.v." stands for "finite variation").

4.2. Default Time and Default Adjusted Return Dynamics. The default time τ admits a predictable intensity $\widehat\mu = (\widehat\mu_t)_{t\ge 0}$ driven by the variance process σ^2. Using the Markov property of σ^2 and the independent and stationary increments property of L, we can establish that $\widehat\mu_t = \widehat\mu(\sigma_{t-})$, where the function $\widehat\mu$ is given by

$$\widehat\mu(x) = \Pi\left(\Big(-\infty, -\frac{1}{x}\Big]\right) = \int_{-\infty}^{-1/x} \Pi(\mathrm{d}y), \;\; x > 0. \tag{4.11}$$

Then the process $N = (N_t)_{t\geq 0}$ defined by

$$N_t = \mathbf{1}_{\{\tau\leq t\}} - \int_0^{t\wedge\tau} \widehat{\mu}(\sigma_{u-})\,\mathrm{d}u$$

is a martingale. The unconditional probability $\mathrm{PD} = (\mathrm{PD}_t)_{t\geq 0}$ of default prior to time t can then be calculated as:

$$\mathrm{PD}_t = P(\tau \leq t) = 1 - \mathbb{E}\left[\exp\left(-\int_0^t \widehat{\mu}(\sigma_u)\,\mathrm{d}u\right)\right],\ t\geq 0\,. \tag{4.12}$$

We now turn to the effect of the default on the dynamics of the driving Lévy process L.

Theorem 4.1. *The bivariate process (S,σ^2) is a Markov process and the stochastic differential of S is given by*

$$\mathrm{d}S_t = S_{t-}\,\left[r + \lambda(\sigma_{t-})\,\sigma_{t-}\right]\,\mathrm{d}t + S_{t-}\,\sigma_{t-}\,\mathrm{d}\widehat{L}_t\,,$$

where $\widehat{L}$ is the stopped version of L with default adjustment

$$\widehat{L}_t = L_t^\tau + \mathbf{1}_{\{t\geq\tau\}}(-\Delta L_\tau - 1/\sigma_{\tau-}),\ t\geq 0.$$

With $\widehat{\lambda}$ defined by

$$\widehat{\lambda}(x) = -\int_{-\infty}^{-1/x}\left(y+\frac{1}{x}\right)\Pi(\mathrm{d}y),\ x>0,$$

the compensated version $(\widehat{L}_t - \int_0^{t\wedge\tau}\widehat{\lambda}(\sigma_{u-})\,\mathrm{d}u)_{t\geq 0}$ is a martingale.

Next, define the default adjusted return process $\widehat{R} = (\widehat{R}_t)_{t\geq 0}$ by

$$\widehat{R}_t = \int_0^{t\wedge\tau}(r + \lambda(\sigma_{u-})\,\sigma_{u-})\,\mathrm{d}u + \int_0^t \sigma_{u-}\,\mathrm{d}\widehat{L}_u\,. \tag{4.13}$$

By construction it is clear that $S = S_0\,\mathcal{E}(\widehat{R})$. It follows that the discounted price process $Z = S/B$ is a local martingale if and only if $(\widehat{R}_t - r\,(t\wedge\tau))_{t\geq 0}$ is a local martingale. (Note that the processes $\widehat{R}$ and S are both stopped at τ.) The next theorem states the semimartingale characteristics of $\widehat{R}$ and is useful for identifying martingale measures. In our setting, the characteristics $(B_t^{\widehat{R}}, C_t^{\widehat{R}}, \Pi_t^{\widehat{R}})_{t\geq 0}$ of $\widehat{R}$ can be expressed as functions of σ_{t-}^2 using the Markov property, see also Kallsen and Vesenmayer [27].

Theorem 4.2. *The semimartingale characteristics $(B_t^{\widehat{R}}, C_t^{\widehat{R}}, \Pi_t^{\widehat{R}})$ of $\widehat{R}$ for the truncation function $h(x) = x\,\mathbf{1}_{\{|x|\leq 1\}}$ are for $t\geq 0$ given by*

$$\begin{aligned}
B_t^{\widehat{R}} &= \mathbf{1}_{\{t<\tau\}}\left(r + \sigma_{t-}\left[\lambda(\sigma_{t-}) + \widehat{\lambda}(\sigma_{t-}) - \int_{1/\sigma_{t-}}^\infty x\,\Pi(\mathrm{d}x)\right]\right),\\
C_t^{\widehat{R}} &= \mathbf{1}_{\{t<\tau\}}\,\sigma_{t-}^2\,\sigma^2,\\
\Pi_t^{\widehat{R}}(A) &= \mathbf{1}_{\{t<\tau\}}\int \mathbf{1}_{\{\sigma_{t-}\,x\in(A\cap(-1,\infty))\}}\,\Pi(\mathrm{d}x) + \mathbf{1}_{\{t<\tau\}}\,\mathbf{1}_{\{-1\in A\}}\,\widehat{\mu}(\sigma_t),
\end{aligned}$$

for Borel sets $A \subset \mathbb{R} \setminus \{0\}$.

Under a martingale measure Q, the drift of $\widehat{R}$ has to reduce to

$$B_t^{\widehat{R},Q} = \mathbf{1}_{\{t<\tau\}} \left(r - \sigma_{t-} \int_{1/\sigma_{t-}}^{\infty} x \, \Pi_t^Q(\mathrm{d}x) \right), \, t \geq 0, \tag{4.14}$$

where Π_t^Q is the jump measure of L_t, and the correction results from our choice of truncation function $h(x) = x \, \mathbf{1}_{\{|x| \leq 1\}}$.

Remark 4.3. In the following we adopt the martingale modeling approach. Madan, Carr, and Chang [34] and Panayotov [41] also use this approach in related settings. Formally, the market model can be investigated for arbitrage using the results provided by Delbaen and Schachermayer [9]. Such a thoroughgoing investigation is beyond the scope of this paper.

4.3. The Risk-Neutral Dynamics and Option Pricing. In the following we assume we are given a measure $Q \sim P$ such that L is a Lévy process with characteristic triplet $(\gamma^Q, (\sigma^Q)^2, \Pi^Q)$ and finite fourth moment. Further, assume that $(\sigma_t^2)_{t \geq 0}$ and S follow the dynamics given in (4.3) and (4.8), with potentially altered parameters to ensure no-arbitrage. (Note that we can assume without loss of generality that L is centered to 0 and scaled to have a unit variance rate.)

The Q-dynamics of σ_t^2 are given by the risk-neutral version of (4.8), i.e.:

$$\mathrm{d}\sigma_t^2 = \kappa^Q \, \left((\overline{\sigma}^Q)^2 - \sigma_{t-}^2 \right) \mathrm{d}t + \nu^Q \, \sigma_{t-}^2 \, \mathrm{d}M_t^Q, \; t > 0, \tag{4.15}$$

where κ^Q, $(\overline{\sigma}^Q)^2$, and ν^Q are the potentially adjusted parameters, and M^Q is the bracket process of L centered to 0 and scaled to unit variance rate. With $\widehat{\lambda}^Q$ defined as in Theorem 4.1 by

$$\widehat{\lambda}^Q(x) = -\int_{-\infty}^{-1/x} \left(y + \frac{1}{x} \right) \Pi^Q(\mathrm{d}y),$$

the compensated version $\widehat{L}_t - \int_0^{t \wedge \tau} \widehat{\lambda}^Q(\sigma_{u-}) \, \mathrm{d}u$ is then a Q-martingale. The risk-neutral return process is then given by the risk-neutral version of (4.13), i.e.:

$$\widehat{R}_t = \int_0^{t \wedge \tau} \left(r - \widehat{\lambda}^Q(\sigma_{u-}) \, \sigma_{u-} \right) \, \mathrm{d}u + \int_0^t \sigma_{u-} \, \mathrm{d}\widehat{L}_u \, . \tag{4.16}$$

The expression $\widehat{\lambda}^Q(\sigma_{t-}) \, \sigma_{t-}$ can be conceptualised as the premium for the limited liability option, i.e. the premium paid by equity for protecting it from losses larger than 100%. The stock price process is given by $S = S_0 \, \mathcal{E}(\widehat{R})$.

Under the measure Q, denote by $\pi^Q(\cdot; \chi)$ the price process of a T-claim χ that is suitably integrable, i.e. the random variable χ is $\mathcal{F}_T$-measurable and $\mathbb{E}^Q|\chi| < \infty$. Then π^Q is given by

$$\pi^Q(t; \chi) = e^{-r\,(T-t)} \, \mathbb{E}^Q \left[\chi \,|\, \mathcal{F}_t \right] \, . \tag{4.17}$$

4.4. Variance-Gamma COGARCH. In this section we take the Variance-Gamma (VG) process proposed by Madan and Seneta [35] and Madan, Carr, and Chang [34], and construct the VG-COGARCH model directly under a martingale measure Q, see discussion in the previous section. We examine the model for its suitability to reflect stylized facts, such as volatility clustering, leptokurtosis and skew, and incorporate as a new feature, possible default. We then compute option prices, and, using the implied Black-Scholes volatility, compare the results to those obtained from a corresponding Heston model. Finally, we discuss our stochastic exponential setup in relation to the exponential VG-COGARCH of Panayotov [41], see Remark 4.4 below.

Under the martingale measure Q the VG process L is defined by $L_t = \theta_{VG}\,\Gamma_t + \sigma_{VG}\,W_{\Gamma_t}$ for $t \geq 0$, where Γ is a Gamma process with variance rate ν_{VG} and unit mean rate carrying the market time. W is a standard Brownian motion independent of Γ, $\sigma_{\mathrm{VG}} > 0$ the volatility, and $\theta_{\mathrm{VG}} \in \mathbb{R}$ the drift. The VG process is a pure jump process having characteristic triplet $(\gamma^Q, 0, \Pi^Q)$ with Lévy measure

$$\Pi^Q(\mathrm{d}x) = \frac{\exp(\theta_{\mathrm{VG}}\, x/\sigma_{\mathrm{VG}}^2)}{|x|\,\nu_{\mathrm{VG}}} \exp\left(- \frac{\sqrt{2\,\sigma_{\mathrm{VG}}^2/\nu_{\mathrm{VG}} + \theta_{\mathrm{VG}}^2}}{\sigma_{\mathrm{VG}}^2}\,|x| \right) \mathrm{d}x\,, \tag{4.18}$$

and drift

$$\gamma^Q = \theta_{\mathrm{VG}} - \int_{|x|>1} x\,\Pi^Q(\mathrm{d}x)\,. \tag{4.19}$$

With this parametrisation, the moments of the Lévy measure are

$$\begin{aligned}
\int x\,\Pi^Q(\mathrm{d}x) &= \theta_{\mathrm{VG}}\,,\\
\int x^2\,\Pi^Q(\mathrm{d}x) &= \theta_{\mathrm{VG}}^2\nu_{\mathrm{VG}} + \sigma_{\mathrm{VG}}^2\,,\\
\int x^3\,\Pi^Q(\mathrm{d}x) &= 2\,\theta_{\mathrm{VG}}^3\nu_{\mathrm{VG}}^2 + 3\,\sigma_{\mathrm{VG}}^2\,\theta_{\mathrm{VG}}\,\nu_{\mathrm{VG}}\,,\\
\int x^4\,\Pi^Q(\mathrm{d}x) &= 6\,\theta_{\mathrm{VG}}^4\,\nu_{\mathrm{VG}}^3 + 12\,\sigma_{\mathrm{VG}}^2\,\theta_{\mathrm{VG}}^2\,\nu_{\mathrm{VG}}^2 + 3\,\sigma_{\mathrm{VG}}^4\,\nu_{\mathrm{VG}}\,.
\end{aligned}$$

Using the normalisation $\int x^2\,\Pi^Q(\mathrm{d}x) = 1$, which forces $\sigma_{\mathrm{VG}}^2 < 1$, the third and fourth moments can be written in the form

$$\begin{aligned}
\int x^3\,\Pi^Q(\mathrm{d}x) &= \mathrm{sign}(\theta_{\mathrm{VG}})\,\sqrt{\nu_{\mathrm{VG}}\,(1-\sigma_{\mathrm{VG}}^2)}\,(2+\sigma_{\mathrm{VG}}^2)\,,\\
\int x^4\,\Pi^Q(\mathrm{d}x) &= 3\,\nu_{\mathrm{VG}}\,(2-\sigma_{\mathrm{VG}}^4)\,.
\end{aligned}$$

Then the leverage in (4.10) is obtained as a function of the VG parameter σ_{VG} and the sign of θ_{VG}:

$$\rho = \mathrm{sign}(\theta_{\mathrm{VG}})\,\frac{\sqrt{1-\sigma_{\mathrm{VG}}^2}\,(2+\sigma_{\mathrm{VG}}^2)}{\sqrt{3\,(2-\sigma_{\mathrm{VG}}^4)}}\,.$$

The risk-neutral default intensity $\widehat{\mu}^Q(x)$ can then be derived from (4.11) as:

$$\widehat{\mu}^Q(x) = \frac{1}{\nu_{\mathrm{VG}}}\, E_1\left(\frac{\theta_{\mathrm{VG}} + \sqrt{2\,\sigma_{\mathrm{VG}}^2/\nu_{\mathrm{VG}} + \theta_{\mathrm{VG}}^2}}{\sigma_{\mathrm{VG}}^2\, x}\right),\ x > 0,$$

where $E_1(x) = \int_x^\infty y^{-1}\, e^{-y}\,\mathrm{d}y$ for $x > 0$.

Figure 1 displays the default intensity $\widehat{\mu}^Q$ depending on the volatility $(\sigma_t)_{t\geq 0}$ for three different parameterisations. The structural parameters of the volatility SDE are $\kappa^Q = 1$, $\overline{\sigma}^Q = 0.30$, $\nu^Q = 1$. The first set of VG-parameters is given by $\theta_{\mathrm{VG}} = -1.64$, $\nu_{\mathrm{VG}} = 0.01$, $\sigma_{\mathrm{VG}}^2 = 0.99$, and reproduces a skew of -0.77 and a kurtosis of 7.90 for daily return data as is typically observed for liquidly traded single stocks (see blue/solid). The second parameter set is given by $\theta_{\mathrm{VG}} = -1.62$, $\nu_{\mathrm{VG}} = 0.02$, $\sigma_{\mathrm{VG}}^2 = 0.97$, and reproduces a skew of -1.51 and a kurtosis of 16.54 for daily return data (black/dotted). This parameter set shows more asymmetries and heavier tails and potentially proxies for rather illiquid mid-cap stocks. The third parameter set is given by $\theta_{\mathrm{VG}} = -1.60$, $\nu_{\mathrm{VG}} = 0.03$, $\sigma_{\mathrm{VG}}^2 = 0.96$, and reproduces a skew of -2.22 and a kurtosis of 25.83 for daily return data (red/dashed). As expected, the default intensity $\widehat{\mu}^Q$ is increasing in the volatility and the kurtosis of the returns.

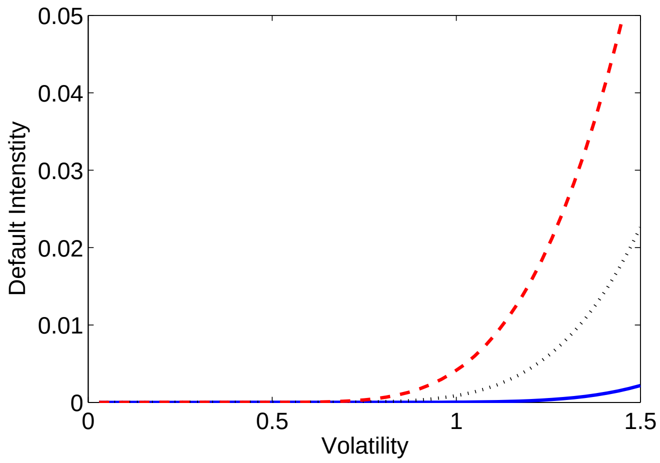

Figure 1. Risk-neutral default intensities $\widehat{\mu}^Q$ for volatility parameters $\kappa^Q = 1$, $\overline{\sigma}^Q = 0.30$, $\nu^Q = 1$, and three different sets of VG parameters: (i) $\theta_{\mathrm{VG}} = -1.64$, $\nu_{\mathrm{VG}} = 0.01$, $\sigma_{\mathrm{VG}}^2 = 0.99$ (blue/solid), (ii) $\theta_{\mathrm{VG}} = -1.62$, $\nu_{\mathrm{VG}} = 0.02$, $\sigma_{\mathrm{VG}}^2 = 0.97$ (black/dotted), or (iii) $\theta_{\mathrm{VG}} = -1.60$, $\nu_{\mathrm{VG}} = 0.03$, $\sigma_{\mathrm{VG}}^2 = 0.96$ (red/dashed).

The compensator $\widehat{\lambda}^Q$ of $\widehat{L}$ can be calculated fairly explicitly as

$$\widehat{\lambda}^Q(x) = \frac{\sigma_{\mathrm{VG}}^2\ \exp\left(-\frac{\theta_{\mathrm{VG}}+\sqrt{2\,\sigma_{\mathrm{VG}}^2/\nu_{\mathrm{VG}}+\theta_{\mathrm{VG}}^2}}{\sigma_{\mathrm{VG}}^2\, x}\right)}{\nu_{\mathrm{VG}}\,(\theta_{\mathrm{VG}} + \sqrt{2\,\sigma_{\mathrm{VG}}^2/\nu_{\mathrm{VG}} + \theta_{\mathrm{VG}}^2})} - \frac{\widehat{\mu}^Q(x)}{x},\ x > 0\,. \tag{4.20}$$

Figure 2 displays the risk-neutral default premium $\widehat{\lambda}^Q(x)$ depending on the volatility for three different parameterisations. The risk-neutral default premium is as expected increasing in the volatility and the kurtosis. The parameterisations are identical to those of Figure 1 discussed above.

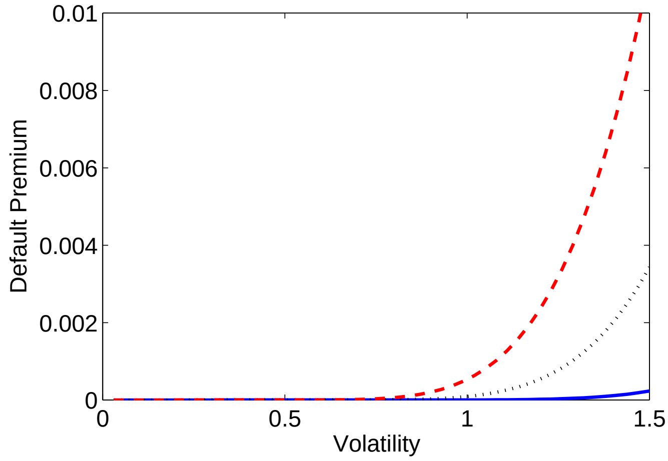

Figure 2. Risk-neutral default premium $\hat{\lambda}^Q(x)\,x$ for volatility parameters $\kappa^Q = 1$, $\overline{\sigma}^Q = 0.30$, $\nu^Q = 1$, and three different sets of VG parameters: (i) $\theta_{\mathrm{VG}} = -1.64$, $\nu_{\mathrm{VG}} = 0.01$, $\sigma^2_{\mathrm{VG}} = 0.99$ (blue/solid), (ii) $\theta_{\mathrm{VG}} = -1.62$, $\nu_{\mathrm{VG}} = 0.02$, $\sigma^2_{\mathrm{VG}} = 0.97$ (black/dotted), or (iii) $\theta_{\mathrm{VG}} = -1.60$, $\nu_{\mathrm{VG}} = 0.03$, $\sigma^2_{\mathrm{VG}} = 0.96$ (red/dashed).

The option pricing model is now completely specified under the martingale measure Q. The driving Lévy process is VG with characterisitc triplet $(\gamma^Q, 0, \Pi^Q)$ defined in (4.18) and (4.19). The volatility dynamics are given according to (4.15) for some κ^Q, $(\overline{\sigma}^Q)^2$, and ν^Q. The risk-neutral default adjusted return process $\widehat{R}$ is then defined according to (4.16) where $\widehat{\lambda}^Q$ is given by (4.20).

Now, we compare the VG-COGARCH to the Heston model. We produce for both models the prices of European call options with varying strike prices and maturities. For the VG-COGARCH we apply Monte-Carlo simulation using a simple Euler discretisation scheme. The Heston call prices are computed by numerical integration of the characteristic function of the log-price process at maturity date. Both prices are then converted to corresponding implied Black-Scholes volatilities. The volatility dynamics is mean reverting around a level of $\overline{\sigma}^Q = 0.30$ with mean reversion rate $\kappa^Q = 1$, for both VG-COGARCH and Heston, and a volatility of volatility parameter of $\nu^Q = 1$ for the VG-COGARCH and $\nu^Q = 0.3$ for Heston, respectively. With this setup we ensure that the volatility dynamics are comparable for both models, see Table 1. The VG parameters are set to $\theta_{\mathrm{VG}} = -1.64$, $\nu_{\mathrm{VG}} = 0.01$, and $\sigma^2_{\mathrm{VG}} = 0.99$ implying a leverage of $\rho = -0.275$ and skewness of -0.77 and kurtosis of 7.90 for the innovations on a daily basis. For the Heston model we have set $\rho = -0.275$ as well to keep the results comparable.

The implied volatility surface for the VG-COGARCH is displayed in Figure 3.

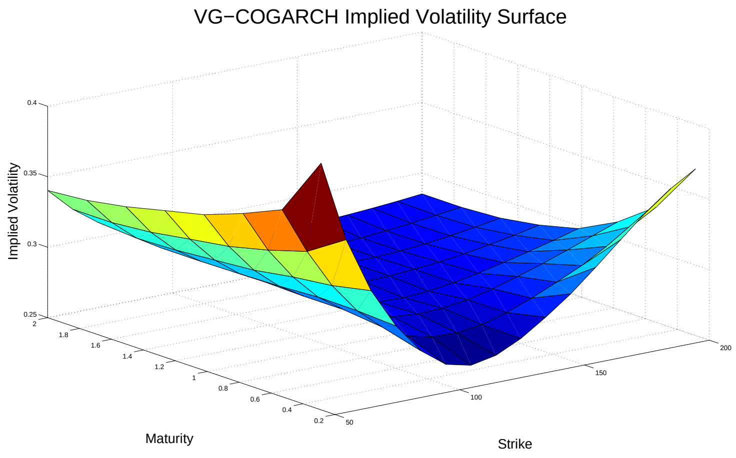

Figure 3. Implied Black-Scholes volatilities from VG-COGARCH call option prices. The risk-neutral parameters are $S_0 = 100$, $r = 0.05$, $\sigma_0^2 = 0.30^2$, $\kappa^Q = 1$, $\overline{\sigma}^Q = 0.30$, $\nu^Q = 1$, $\theta_{\text{VG}} = -1.64$, $\nu_{\text{VG}} = 0.01$, $\sigma^2_{\text{VG}} = 0.99$.

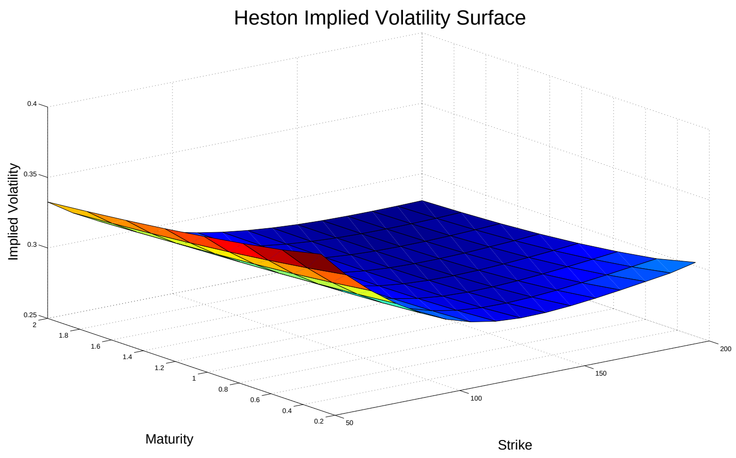

Figure 4. Implied Black-Scholes volatilities from Heston call option prices. The risk-neutral parameters are $S_0 = 100$, $r = 0.05$, $\sigma_0^2 = 0.30^2$, $\kappa^Q = 1$, $\overline{\sigma}^Q = 0.30$, $\nu^Q = 0.30$, $\rho = -0.27$.

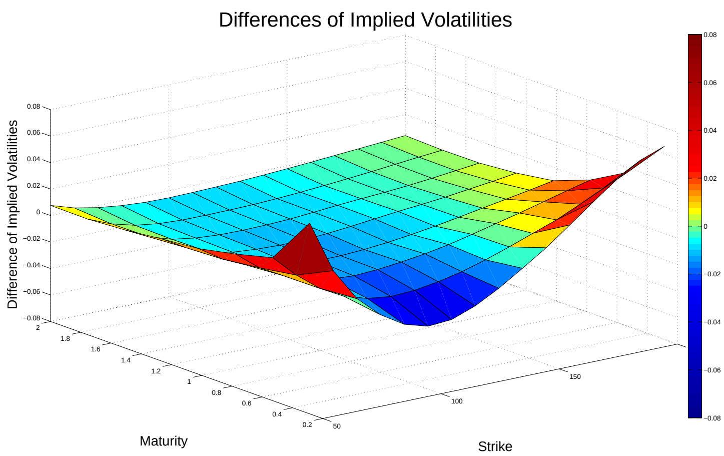

Figure 5. Difference of implied volatilities in Figure 3 and Figure 4.

The jumps and accordingly the high kurtosis lead to rather steep smile patterns for short dated options. At the long end the skewness dominates, and the typical smirk can be observed with declining implied volatilities for increasing strike prices. For the corresponding Heston model the implied volatility surface is graphed in Figure 4. Here, the smile for short dated options is of rather mild extent. This finding is well-known and can be attributed to the continuous price paths inherent in the Heston model. For longer maturities the skewness generated by the negative correlation $\rho = -0.275$ produces an implied volatility smirk approximately of the same extent as observed for VG-COGARCH. A difference plot for both volatilities is given in Figure 5. One may summarise that the VG-COGARCH can produce higher implied volatilities for short maturities deep in-the-money and far out-of-the-money.

Remark 4.4. We conclude this section by mentioning that a similar option pricing procedure for the COGARCH model has also been suggested by Panayotov [41]. In contrast to us he models the risk-neutral dynamics of the log price by a VG-COGARCH process leading to a stock price process

$$\widetilde{S}_t = \widetilde{S}_0 \exp\Big(\int_0^t a_u \mathrm{d}u + \int_0^t \sigma_{u-}\, \mathrm{d}L_u \Big) ,$$

where $(\sigma_t)_{t\geq 0}$ is the COGARCH volatility driven by the VG process L, see Panayotov [41], Eq. (3.3.1). The expression $\int_0^t a_u \mathrm{d}u$ is a convexity correction which guarantees that the stock price has the proper risk-neutral expectation. According to (3.3.4) in Panayotov [41] the density of the correction a can be computed as follows

$$a_t = r - \int_{-\infty}^{\infty} \Big(e^{\sigma_t\, x} - 1 \Big)\, \Pi^Q(\mathrm{d}x),\; t \geq 0\, .$$

The log price process can be derived from this, and, using the fact that, together with the volatility process, it is jointly Markovian, the option price is calculated by numerically solving a PIDE. Possibility of default is not included in his model.

5. Statistical Estimation of COGARCH

We present two different estimation procedures. The first is a simple method of moments estimation, which works only for equally spaced data. The second method is more sophisticated and handles unequally spaced data. It needs some preliminary results concerning the pathwise approximation of Lévy processes which we outline in this section. Throughout this section we assume that the driving Lévy process has no Gaussian part, i.e. that $\sigma^2 = 0$.

5.1. A Method of Moments Estimation. For practical purposes, we need to discretise the continuous-time COGARCH onto a discrete grid over a finite time interval, and with a finite state space. Assume first that our data are given as described in (3.9). The goal of this section is to estimate the model parameters β, η, φ. Moreover, we shall present a simple estimate of the volatility.

5.1.1. Identifiability of the model parameters. We aim at estimation of the model parameters (β, η, φ) from a sample of equally spaced returns by matching empirical autocorrelation function and moments to their theoretical counterparts given in Proposition 3.1. The next result shows that the parameters are identifiable by this estimation procedure for driving Lévy processes L as in Proposition 3.1. We assume throughout that $\mathbb{E}(L_1) = 0$ and $\mathbb{E}(L_1^2) = 1$. For the sake of simplicity we set $r = 1$.

Theorem 5.1. *Suppose $(L_t)_{t\geq 0}$ is a Lévy process such that $\mathbb{E}(L_1) = 0$, $\mathbb{E}(L_1^2) = 1$, $\mathbb{E}(L_1^4) < \infty$ and $\int_{\mathbb{R}} x^3\, \Pi(dx) = 0$. Assume also that $\Psi(2) < 0$, and denote by $(G_i^{(1)})_{i\in\mathbb{N}}$ the stationary increment process of the* COGARCH(1,1) *process with parameters $\beta, \eta, \varphi > 0$. Let $\mu, \gamma(0), k, p > 0$ be constants such that*

$$\begin{aligned} \mathbb{E}((G_i^{(1)})^2) &= \mu, \\ \mathrm{Var}((G_i^{(1)})^2) &= \gamma(0), \\ \rho(h) = \mathrm{corr}((G_i^{(1)})^2, (G_{i+h}^{(1)})^2) &= k e^{-hp}, \quad h \in \mathbb{N}. \end{aligned}$$

Define

$$\begin{aligned} M_1 &:= \gamma(0) - 2\mu^2 - 6\frac{1-p-e^{-p}}{(1-e^p)(1-e^{-p})}\, k\,\gamma(0), \\ M_2 &:= \frac{2k\gamma(0)p}{M_1(e^p-1)(1-e^{-p})}. \end{aligned}$$

Then $M_1, M_2 > 0$, and the parameters β, η, φ are uniquely determined by $\mu, \gamma(0), k$ and p and are given by the formulas

$$\beta = p\,\mu\,, \tag{5.1}$$

$$\varphi = p\sqrt{1+M_2} - p, \tag{5.2}$$

$$\eta = p\sqrt{1+M_2}\,1 + p = p\,. \tag{5.3}$$

We conclude from (5.1)–(5.3) that our model parameter vector (β, η, φ) is a continuous function of the first two moments $\mu, \gamma(0)$ and the parameters of the autocorrelation function p and k. Hence, by continuity, consistency of the moments will immediately imply consistency of the corresponding plug-in estimates for (β, η, φ).

5.1.2. The estimation algorithm. The parameters are estimated under the following assumptions:

(H1) We have equally spaced observations G_i, $i = 0, \dots, n$, on the integrated COGARCH as defined and parameterised in (3.1) and (3.2), assumed to be in its stationary regime. This gives return data $G_i^{(1)} = G_i - G_{i-1}$, $i = 1, \dots, n$.

(H2) $\mathbb{E}(L_1) = 0$ and $\mathbb{E}(L_1^2) = 1$, i.e. $(\sigma_t^2)_{t\ge 0}$ can be interpreted as the volatility.

(H3) The driving Lévy process has no Gaussian part.

(H4) $\int_{\mathbb{R}} x^3\,\Pi(dx) = 0$, $\mathbb{E}(L_1^4) < \infty$ and $\Psi(2) < 0$.

We proceed as follows.

(1) Calculate the moment estimator $\widehat{\mu}_n$ of μ as

$$\widehat{\mu}_n := \frac{1}{n}\sum_{i=1}^{n}(G_i^{(1)})^2,$$

and for fixed $d \ge 2$ the empirical autocovariances $\widehat{\boldsymbol{\gamma}}_n := (\widehat{\gamma}_n(0), \widehat{\gamma}_n(1), \dots, \widehat{\gamma}_n(d))^T$ as

$$\widehat{\gamma}_n(h) := \frac{1}{n}\sum_{i=1}^{n-h}\left((G_{i+h}^{(1)})^2 - \widehat{\mu}_n\right)\left((G_i^{(1)})^2 - \widehat{\mu}_n\right), \quad h = 0, \dots, d\,.$$

(2) Compute the empirical autocorrelations $\widehat{\boldsymbol{\rho}}_n := (\widehat{\gamma}_n(1)/\widehat{\gamma}_n(0), \dots, \widehat{\gamma}_n(d)/\widehat{\gamma}_n(0))^T$.

(3) For fixed $d \ge 2$ define the mapping $H : \mathbb{R}_+^{d+2} \to \mathbb{R}$ by

$$H(\widehat{\boldsymbol{\rho}}_n, \boldsymbol{\theta}) := \sum_{h=1}^{d}(\log(\widehat{\rho}_n(h)) - \log k + ph)^2\,.$$

Compute the least squares estimator[4]

$$\widehat{\boldsymbol{\theta}}_n := \operatorname{argmin}_{\boldsymbol{\theta}\in\mathbb{R}_+^2} H(\widehat{\boldsymbol{\rho}}_n, \boldsymbol{\theta}) \,. \tag{5.4}$$

(4) Define the mapping $J : \mathbb{R}_+^4 \to [0,\infty)^3$ by

$$J(\mu, \gamma(0), \boldsymbol{\theta}) \quad := \quad \begin{cases} (p\mu, p\sqrt{1+M_2} - p, p\sqrt{1+M_2} + p) & \text{if} \quad p, M_2 > 0\,, \\ (0,0,0) & \text{otherwise,} \end{cases}$$

where M_2 is defined as in (5.1). Finally, compute the estimator

$$\widehat{\boldsymbol{\vartheta}}_n = J(\widehat{\mu}_n, \widehat{\gamma}_n(0), \widehat{\boldsymbol{\theta}}_n) \,.$$

In Haug et al. [24] asymptotic normality of the estimated parameter vector was proved. This is essentially a consequence of the geometric ergodicity of the returns process $(G_i^{(1)})_{i\in\mathbb{N}}$.

To conclude this section we mention that Müller [38] developed an MCMC estimation procedure for the COGARCH(1,1) model, which works also for irregularly spaced observations. The approach is, however, restricted to driving processes L of finite variation. Alternatively, Fasen [17] presents results on the non-parametric estimation of the autocovariance function of the volatility process and the COGARCH process by invoking point process methods. In the next section, we outline a more sophisticated way of dealing with unequally spaced data. It applies some results concerning the pathwise approximation of Lévy processes.

5.2. The "First Jump" Approximation for a Lévy Process. In this section we review a "first jump" approximation to the underlying Lévy process which preserves certain crucial features of the process.

Suppose again that the Lévy process $(L_t)_{t\geq 0}$ has characteristic triplet of the form $(\gamma, 0, \Pi)$, where $\gamma \in \mathbb{R}$ and Π is the Lévy measure. As usual, denote the jumps of L_t by $\Delta L_t = L_t - L_{t-}$ for $t \geq 0$ (with $L_{0-} = 0$), and let

$$\overline{\Pi}(x) = \Pi((x,\infty)) + \Pi((-\infty,-x]),\; x > 0, \tag{5.5}$$

denote the tail of $\Pi(\cdot)$.

We wish to approximate L on a finite time interval $[0,T]$, $0 < T < \infty$, partitioned into N_n not necessarily equally spaced intervals. Let $(N_n)_{n\in\mathbb{N}}$ be an increasing sequence of integers diverging to infinity as $n \to \infty$. For each $n \in \mathbb{N}$, form a deterministic partition $0 = t_0(n) < t_1(n) < \cdots < t_{N_n}(n) = T$ of $[0,T]$. In Maller and Szimayer [46], two approximating processes to L are constructed.

[4] We note the known robustness issues associated with least squares estimators. There is no difficulty in substituting for $\widehat{\boldsymbol{\theta}}_n$ a more robust estimator, for instance, by replacing the $\mathcal{L}^2$-norm by the $\mathcal{L}^1$-norm, or invoking a weighted Huber estimator.

The first approximation, $\overline{L}_t(n)$ for $n \in \mathbb{N}$ is formed by taking the first jump, if one occurs, of L_t in each time subinterval $(t_{j-1}(n), t_j(n)]$, $j = 1, 2, \ldots, N_n$, where the jump sizes are bounded away from 0, then discretizing ("binning") these jumps to get an approximating process which takes only a finite number of values on a finite state space. The state space does not include 0, as we must avoid the possible singularity in Π at 0. If no jump occurs in a subinterval, $\overline{L}_t(n)$ remains constant in that subinterval.

A second approximating process, $L_t(n)$, $n \in \mathbb{N}$, is then taken as the discrete skeleton of $\overline{L}_t(n)$ on the time grid $(t_j(n))_{j=0,1,\ldots,N_n}$.

The time and space intervals are allowed to shrink and the state space to expand at appropriate rates, so as to get convergence of $\overline{L}_t(n)$ and $L_t(n)$ to L_t, as $n \to \infty$, in various modes.

To see how this works, take two sequences of real numbers $(m_n)_{n\in\mathbb{N}}$ and $(M_n)_{n\in\mathbb{N}}$, satisfying $1 > m_n \downarrow 0$ and $1 < M_n \uparrow \infty$, as $n \to \infty$. The first approximating process, $\overline{L}_t(n)$, takes discrete values in the set

$$J(n) = [-M_n, -m_n) \cup (m_n, M_n],\ n \in \mathbb{N}.$$

To construct it, let

$$\tau_j(n) := \inf\{t : t_{j-1}(n) < t \le t_j(n); \Delta L_t \in J(n)\}, \quad \text{for } 1 \le j \le N_n,$$

(where the infimum over the empty set is defined as ∞) be the time of the first jump of L with magnitude in $(m_n,\, M_n]$ in interval j. Then decompose L_t as

$$L_t = \gamma_n\, t + L_t^{(1)}(n) + L_t^{(2)}(n) + L_t^{(3)}(n)\,, \quad \text{for } 0 \le t \le T\,, \tag{5.6}$$

where for all $n \ge 1$ and $0 \le t \le T$:

$$\begin{aligned} L_t^{(1)}(n) &= \text{a.s.}\lim_{\varepsilon\downarrow 0}\left(\sum_{0<s\le t} \Delta L_s\, \mathbf{1}_{\{\varepsilon<|\Delta L_s|\le m_n\}} - t\int_{\varepsilon<|x|\le m_n} x\,\Pi(\mathrm{d}x)\right), \\ L_t^{(2)}(n) &= \sum_{0<s\le t} \Delta L_s\, \mathbf{1}_{\{M_n<|\Delta L_s|\}}, \quad L_t^{(3)}(n) = \sum_{0<s\le t} \Delta L_s\, \mathbf{1}_{\{m_n<|\Delta L_s|\le M_n\}}, \end{aligned}$$

and

$$\gamma_n = \gamma - \int_{m_n<|x|\le 1} x\,\Pi(\mathrm{d}x).$$

Decomposition (5.6) is a variant of the Lévy-Ito decomposition (Sato [44], Theorem 19.2, p. 120), in which, for each n, $L_t^{(1)}(n)$ is a compensated "small jump" martingale, and $L_t^{(2)}(n)$ and $L_t^{(3)}(n)$ might be thought of as "large jumps" and "medium jumps", respectively.

With no assumptions on L, Szimayer and Maller [46] show that, for $j = 1, 2$, $\lim_{n\to\infty} \sup_{0\le t\le T} L_t^{(j)}(n) = 0$ a.s. $L_t^{(3)}(n)$ can be further decomposed as follows:

$$L_t^{(3)}(n) = L_t^{(3,1)}(n) + L_t^{(3,2)}(n), \tag{5.7}$$

where

$$L_t^{(3,2)}(n) = \sum_{j=1}^{N_n} \mathbf{1}_{\{\tau_j(n) \le t\}} \Delta L^{(3)}_{\tau_j(n)}. \tag{5.8}$$

Thus $L_t^{(3,2)}(n)$ is the sum of the sizes of the first jump of L_t in each subinterval whose magnitude is in $(m_n, M_n]$, where such jumps occur, while $L_t^{(3,1)}(n)$ collects, over all subintervals, the sizes of those jumps with magnitudes in $(m_n, M_n]$ (except for the first jump), provided at least two such jumps occur in a subinterval.

Since we allow for the possibility that L has "infinite activity", that is, that $\Pi(\mathbb{R} \setminus \{0\}) = \infty$, we need a restriction on how fast m_n may tend to the possible singularity of Π at 0, by comparison with the speed at which the time mesh shrinks. With appropriate assumptions, $\lim_{n\to\infty} \sup_{0\le t\le T} |L_t^{(3,1)}(n)| = 0$ in probability, in $\mathcal{L}_1$, or, alternatively, in the almost sure sense. This leaves $L^{(3,2)}(n)$ as the predominant component, asymptotically, of L, and the penultimate step is to approximate it by a process $\overline{L}(n)$ that lives on a finite state space. So we discretize the state space $J(n)$ with a grid of mesh size $\Delta(n) > 0$, where $\Delta(n) \searrow 0$ as $n \to \infty$, and set

$$\overline{L}_t(n) = \gamma_n t + \sum_{j=1}^{N_n} \mathbf{1}_{\{\tau_j(n) \le t\}} \left\lfloor \frac{\Delta L^{(3)}_{\tau_j(n)}}{\Delta(n)} \right\rfloor \Delta(n)\,. \tag{5.9}$$

(The symbol $\lfloor x \rfloor$ denotes the integer part of $x \in \mathbb{R}$). Again under certain conditions, the difference between $L^{(3,2)}(n)$ and $\overline{L}(n)$ disappears, asymptotically, in the $\mathcal{L}_1$ or almost sure sense, uniformly in $0 \le t \le T$. Thus $\overline{L}(n)$ approximates L, in the sense that the distance between them as measured by the supremum metric tends to 0 in $\mathcal{L}_1$ or almost surely, in our setup.

The second approximation, $L(n)$, is obtained by evaluating $\overline{L}(n)$ on the same discrete time grid as we have used so far. Thus $L(n)$ is the piecewise constant process defined by

$$L_t(n) = \overline{L}_{t_{j-1}(n)}(n), \text{ when } t_{j-1}(n) \le t < t_j(n),\ j = 1, 2, \ldots, N_n, \tag{5.10}$$

and with $L_T(n) = \overline{L}_T(n)$. Because the original jumps are displaced in time in $L(n)$, we no longer expect convergence to L in the supremum metric. Instead, we get that $\lim_{n\to\infty} \rho(L(n), L) = 0$, where $\rho(\cdot,\cdot)$ denotes the Skorokhod J_1 distance in $\mathbb{D}[0,T]$. The processes $L(n)$ approximate L, *pointwise*, in probability, but not uniformly in $0 \le t \le T$. However, the convergence in probability in the Skorokhod topology suffices for certain applications that we discuss later.

Now we state the theorems from [46], which give the convergence of $\overline{L}_t(n)$ and $L_t(n)$ to L_t. Recall from (5.5) that $\overline{\Pi}$ denotes the tail of the Lévy measure of L_t. Let

$$\Delta t(n) := \max_{1\le j\le n} \left(t_j(n) - t_{j-1}(n)\right).$$

The main result for $\overline{L}_t(n)$ is:

Theorem 5.2. *(a) Suppose*

$$\lim_{n\to\infty} \Delta t(n)\,\overline{\Pi}^2(m_n) = 0 \text{ and } \lim_{n\to\infty} \Delta(n)\,\overline{\Pi}(m_n) = 0. \tag{5.11}$$

Then

$$\sup_{0\le t\le T} \left|\overline{L}_t(n) - L_t\right| \stackrel{P}{\to} 0, \text{ as } n\to\infty. \tag{5.12}$$

Next we consider the second approximating process, $L_t(n)$, as defined in (5.10). With a view to applications, we need the following property. The processes $(L_t(n))_{n\in\mathbb{N}}$ are said to satisfy *Aldous' criterion for tightness if:*

$$\forall\varepsilon > 0 : \lim_{\delta\searrow 0} \limsup_{n\to\infty} \sup_{\sigma,\tau\in\mathcal{S}_{0,T}(n),\, \sigma\le\tau\le\sigma+\delta} \mathbb{P}\left(|L_\tau(n) - L_\sigma(n)| \ge \varepsilon\right) = 0\,, \tag{5.13}$$

where $\mathcal{S}_{t,T}(n)$ is the set of $\mathbb{F}^{L(n)}$-stopping times taking values in $[t,T]$, for $0\le t\le T$. Let $\mathbb{D}[0,T]$ be the space of càdlàg real-valued functions on $[0,T]$ and $\rho(\cdot,\cdot)$ the Skorokhod J_1 distance between two processes in $\mathbb{D}[0,T]$.

Theorem 5.3. *Assume that Condition* (5.11) *of Theorem 5.2 holds. Then:*
(i) $\rho(L(n),L) \stackrel{P}{\to} 0$ *as* $n\to\infty$*;*
(ii) the sequence $(L_t(n))_{n\in\mathbb{N}}$ *satisfies Aldous' criterion for tightness.*

We conclude this section with some comments on the filtrations. Let $\mathbb{F}^L$, $\mathbb{F}^{\overline{L}(n)}$ and $\mathbb{F}^{L(n)}$ be the natural filtrations generated by the processes $(L_t)_{t\ge 0}$, $(\overline{L}_t(n))_{t\ge 0}$ and $(L_t(n))_{t\ge 0}$, respectively. Our construction clearly gives inclusion of the filtrations, that is, for each $n\ge 1$

$$\mathbb{F}^{L(n)} \subseteq \mathbb{F}^{\overline{L}(n)} \subseteq \mathbb{F}^L, \tag{5.14}$$

so, having demonstrated convergence of the approximating processes, we will have sufficient structure to prove convergence in some optimal stopping problems using recent results of Coquet and Toldo [7]. More discussion and possible applications of this can be found in Maller and Szimayer [46].

5.3. A Discrete Approximation to the COGARCH. In this section we show how to approximate a COGARCH pair $(G_t,\sigma_t)_{t\ge 0}$ with an embedded sequence of discrete time GARCH pairs, $(G_n(t),\sigma_n(t))_{t\ge 0}$, using the first jump technology developed in Section 5.2. The discrete approximating sequence, after appropriate rescaling, converges to the continuous time model in probability, in the Skorokhod metric, as the discrete approximating grid grows finer. This construction opens the way to using, for the COGARCH, similar statistical techniques to those already worked out for GARCH models, and useful applications can be made to options pricing, and to the modelling of irregularly spaced time series data.

For these kinds of applications L is usually assumed to have finite variance and mean 0, as we will do throughout this section.

Thus, we take as given the continuous time COGARCH pair $(G_t,\sigma_t)_{t\ge 0}$ defined in (3.1) and (3.2), and form a discrete approximating sequence as follows. Fix

$T > 0$, and take deterministic sequences $(N_n)_{n\in\mathbb{N}}$ with $\lim_{n\to\infty} N_n = \infty$ and $0 = t_0(n) < t_1(n) < \ldots < t_{N_n}(n) = T$, and, for each $n \in \mathbb{N}$, divide $[0,T]$ into N_n subintervals of length $\Delta t_i(n) := t_i(n) - t_{i-1}(n)$, for $i = 1, 2, \ldots, N_n$. Assume $\Delta t(n) := \max_{i=1,\ldots,N_n} \Delta t_i(n) \to 0$ as $n \to \infty$, and define, for each $n \in \mathbb{N}$, a discrete time process $(G_{i,n})_{i=1,\ldots,N_n}$ satisfying

$$G_{i,n} = G_{i-1,n} + \sigma_{i-1,n}\sqrt{\Delta t_i(n)}\varepsilon_{i,n},\ i = 1, 2, \ldots, N_n, \tag{5.15}$$

where $G_{0,n} = G_0 := 0$, and the variance $\sigma^2_{i,n}$ follows the recursion

$$\sigma^2_{i,n} = \beta\Delta t_i(n) + \left(1 + \varphi\Delta t_i(n)\varepsilon^2_{i,n}\right) e^{-\eta\Delta t_i(n)}\sigma^2_{i-1,n},\ i = 1, 2, \ldots, N_n. \tag{5.16}$$

Here the innovations $(\varepsilon_{i,n})_{i=1,\ldots,N_n}$, $n \in \mathbb{N}$, are constructed using the "first jump" approximation outlined in Section 5.2. Since we assume a finite variance for L, we need only a single sequence $1 \geq m_n \downarrow 0$ bounding the jumps of L away from 0. We assume it satisfies $\lim_{n\to\infty} \Delta t(n)\overline{\Pi}^2_L(m_n) = 0$. Such a sequence always exists, as $\lim_{x\downarrow 0} x^2\,\overline{\Pi}_L(x) = 0$. Fix $n \geq 1$ and define stopping times $\tau_{i,n}$ by

$$\tau_{i,n} = \inf\left\{t \in [t_{i-1}(n), t_i(n)) : |\Delta L(t)| \geq m_n\right\}, \quad i = 1, \ldots, N_n. \tag{5.17}$$

Thus $\tau_{i,n}$ is the time of the first jump of L in the ith interval whose magnitude exceeds m_n, if such a jump occurs.

By the strong Markov property, $\left(\mathbf{1}_{\{\tau_{i,n}<\infty\}}\Delta L(\tau_{i,n})\right)_{i=1,\ldots,N_n}$ is for each $n \in \mathbb{N}$ a sequence of independent rvs, with distribution specified by:

$$\frac{\Pi(\mathrm{d}x)\mathbf{1}_{\{|x|>m_n\}}}{\overline{\Pi}(m_n)}\left(1 - e^{-\Delta t_i(n)\,\overline{\Pi}(m_n)}\right),\ x \in \mathbb{R}\setminus\{0\},\ i = 1, 2, \ldots, N_n, \tag{5.18}$$

and with mass $e^{-\Delta t_i(n)\overline{\Pi}(m_n)}$ at 0. These rvs have finite mean, $\nu_i(n)$, and variance, $\xi_i(n)$, say. The innovations series $(\varepsilon_{i,n})_{i=1,\ldots,N_n}$ required for (5.15) is now defined by

$$\varepsilon_{i,n} = \frac{\mathbf{1}_{\{\tau_{i,n}<\infty\}}\,\Delta L(\tau_{i,n}) - \nu_i(n)}{\xi_i(n)}, \quad i = 1, 2, \ldots, N_n. \tag{5.19}$$

For each $n \in \mathbb{N}$, the $\varepsilon_{i,n}$ are independent rvs with $\mathbb{E}\varepsilon_{1,n} = 0$ and $\mathrm{Var}(\varepsilon_{1,n}) = 1$. Finally, in (5.16), we take $\sigma^2_{0,n} = \sigma^2_0$, independent of the $\varepsilon_{i,n}$.

Remark 5.4. Equations (5.15) and (5.16) specify a GARCH(1,1)-type recursion in the following sense. In the ordinary discrete time GARCH(1,1) series, the volatility sequence satisfies (2.2), viz.,

$$\sigma^2_i = \beta + \left(1 + (\phi/\delta)\varepsilon^2_{i-1}\right)\delta\sigma^2_{i-1}. \tag{5.20}$$

When the time grid is equally spaced, so $\Delta t_i(n) = \Delta t(n)$, $i = 1, 2, \ldots, N_n$, (5.16) is equivalent to (5.20), after rescaling by $\Delta t(n)$ and a reparametrisation from (β, φ, η) to (β, φ, δ), and (5.15) becomes a rescaled GARCH equation for the differenced sequence $G_{i,n} - G_{i-1,n}$. More generally, with an unequally spaced grid, if the series are scaled as in (5.15) and (5.16), convergence to the COGARCH is obtained, as follows.

Embed the discrete time processes $G_{\cdot,n}$ and $\sigma^2_{\cdot,n}$ into continuous time versions G_n and σ^2_n defined for $0 \le t \le T$ by

$$G_n(t) := G_{i,n} \quad \text{and} \quad \sigma^2_n(t) := \sigma^2_{i,n}, \quad \text{when } t \in [t_{i-1}(n), t_i(n)), \tag{5.21}$$

with $G_n(0) = 0$. The processes G_n and σ_n are in $\mathbb{D}[0, T]$. The next result is proved in Theorem 2.1 of Maller, Müller and Szimayer [36].

Theorem 5.5. *In the above setup, the Skorokhod distance between the processes* (G, σ^2) *defined by* (3.1) *and* (3.2)*, and the discretised, piecewise constant processes* $(G_n, \sigma^2_n)_{n\ge1}$ *defined by* (5.21)*, converges in probability to 0 as* $n \to \infty$*; that is,*

$$\rho\left((G_n, \sigma^2_n), (G, \sigma^2)\right) \xrightarrow{\mathbb{P}} 0, \ \text{ as } n \to \infty. \tag{5.22}$$

Consequently, we also have convergence in distribution in $\mathbb{D}[0, T] \times \mathbb{D}[0, T]$*:* $(G_n, \sigma^2_n) \xrightarrow{D} (G, \sigma^2)$*, as* $n \to \infty$*.*

Remark 5.6. Kallsen and Vesenmayer [27] derive the infinitesimal generator of the bivariate Markov process representation of the COGARCH model and show that any COGARCH process can be represented as the limit in law of a sequence of GARCH(1,1) processes. The result of Theorem 5.5 is stronger in that it gives convergence to the continuous-time model in a strong sense (in probability, in the Skorokhod metric), as the discrete approximating grid grows finer. Whereas the diffusion limit in law established by Nelson [40] occurs from GARCH by aggregating its innovations, the COGARCH limit arising in Kallsen and Vesenmayer [27] and Maller et al. [36] both occur when the innovations are randomly thinned.

5.4. GARCH Analysis of Irregularly Spaced Data. Maller, Müller and Szimayer [36] apply the discrete approximation of the continuous time GARCH process to develop a method of fitting the model to unequally spaced times series data, using the methodology worked out for the discrete time GARCH.

5.4.1. The estimation algorithm. The parameters are estimated under the following assumptions:

(H1) Suppose given observations G_{t_i}, $0 = t_0 < t_1 < \ldots < t_N = T$, on the integrated COGARCH as defined and parameterised in (3.1) and (3.2), assumed to be in its stationary regime.

(H2) The (t_i) are assumed fixed (non-random) time points.

(H3) $\mathbb{E}L(1) = 0$ and $\mathbb{E}L^2(1) = 1$; i.e. σ^2 can be interpreted as the volatility.

(H4) The driving Lévy process has no Gaussian part.

Then we proceed as follows.

(1) Let $Y_i = G_{t_i} - G_{t_{i-1}}$ denote the observed increments and put $\Delta t_i := t_i - t_{i-1}$. Then from (3.1) we can write

$$Y_i = \int_{t_{i-1}}^{t_i} \sigma_{s-} \mathrm{d}L(s). \tag{5.23}$$

(2) We can use a pseudo-maximum likelihood (PML) method to estimate the parameters (β, η, φ) from the observed $Y_1, Y_2, \dots, Y_N$. The pseudo-likelihood function can be derived as follows. Because $(\sigma_t)_{t\geq 0}$ is Markovian, Y_i is conditionally independent of $Y_{i-1}, Y_{i-2}, \dots$, given $\mathcal{F}_{t_{i-1}}$. We have $\mathbb{E}(Y_i \mid \mathcal{F}_{t_{i-1}}) = 0$ for the conditional expectation of Y_i, and, for the conditional variance,

$$\rho_i^2 := \mathbb{E}(Y_i^2 \mid \mathcal{F}_{t_{i-1}}) = \left(\sigma_{t_{i-1}}^2 - \frac{\beta}{\eta - \varphi}\right)\left(\frac{e^{(\eta-\varphi)\Delta t_i} - 1}{\eta - \varphi}\right) + \frac{\beta \Delta t_i}{\eta - \varphi}. \tag{5.24}$$

Eq. (5.24) follows from the calculation in the third display on p. 618 of Klüppelberg et al. [30]. To ensure stationarity, we take $\mathbb{E}\sigma_0^2 = \beta/(\eta - \varphi)$, with $\eta > \varphi$, in that formula.

(3) Applying the PML method, then, we assume that the Y_i are conditionally $N(0, \rho_i^2)$, and use recursive conditioning to write a pseudo-log-likelihood function for the observations $Y_1, Y_2, \dots, Y_N$ as

$$\mathcal{L}_N = \mathcal{L}_N(\beta, \varphi, \eta) = -\frac{1}{2}\sum_{i=1}^N \left(\frac{Y_i^2}{\rho_i^2}\right) - \frac{1}{2}\sum_{i=1}^N \log(\rho_i^2) - \frac{N}{2}\log(2\pi). \tag{5.25}$$

(4) We must substitute in (5.25) a calculable quantity for ρ_i^2, hence we need such for $\sigma_{t_{i-1}}^2$ in (5.24). For this, we discretise the continuous time volatility process just as was done in Theorem 5.5. Thus, (5.16) reads, in the present notation,

$$\sigma_i^2 = \beta \Delta t_i + e^{-\eta \Delta t_i}\sigma_{i-1}^2 + \varphi e^{-\eta \Delta t_i} Y_i^2. \tag{5.26}$$

(5) Finally, note that (5.26) is a GARCH-type recursion, so, after substituting σ_{i-1}^2 for $\sigma_{t_{i-1})}^2$ in (5.24), and the resulting modified ρ_i^2 in (5.25), we can think of (5.25) as the pseudo-log-likelihood function for fitting a GARCH model to the unequally spaced series.

The recursion in (5.26) is easily programmed, and, taking as starting value for σ_0^2 the stationary value $\beta/(\eta - \varphi)$, we can maximise the function $\mathcal{L}_N$ to get PMLEs of (β, η, φ). The small sample behaviour of these estimates are investigated in a simulation study in Durand, Maller and Müller [14]. Moreover, Müller, Maller and Durand [39] and Durand, Maller and Müller [14] apply this method to various financial data sets.

Acknowledgements. The authors are grateful to Jan Kallsen for useful discussions regarding the COGARCH option pricing setup.

References

[1] O.E. Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Finance & Stochastics*, 2:41–68, 1998.

[2] D. S. Bates. Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options. *Rev. Fin. Stud.*, 9:69–107, 1996.

[3] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Pol. Econ.*, 87:637–659, 1973.

[4] T. Bollerslev. Generalised autoregressive conditionally heteroscedasticity. *J. Econometrics*, 31:307–327, 1986.

[5] P. J. Brockwell, E. Chadraa, and A. M. Lindner. Continuous time GARCH processes of higher order. *Ann. Appl. Probab.*, 16:790–826, 2006.

[6] L.D. Brown, Y Wang, and L.H. Zhao. On the statistical equivalence at suitable frequencies of GARCH and stochastic volatility models with the corresponding diffusion model. *Statistica Sinica*, 13:993–1013, 2003.

[7] F. Coquet and S. Toldo. Convergence of values in optimal stopping and convergence of optimal stopping times. *El. J. Probab.*, 12(8):207–228, 2007.

[8] V. Corradi. Reconsidering the continuous time limit of the GARCH(1,1) process. *J. Econ.*, 96:145–153, 2000.

[9] F. Delbaen and W. Schachermayer. The fundamental theorem of asset pricing for unbounded stochastic processes. *Math. Annalen*, 312:215–250, 1989.

[10] P. Diaconis and D. Freedman. Iterated random functions. *SIAM Rev.*, 41(1):45–76, 1999.

[11] O.A. Dickey and W.A. Fuller. Distribution for the estimates for auto-regressive time series with a unit root. *J. Amer. Statist. Assoc.*, 74:427–431, 1979.

[12] J. L. Doob. The elementary Gaussian processes. *Ann. Math. Stat.*, 15:229–282, 1944.

[13] J.C. Duan. Augmented GARCH(p,q) process and its diffusion limit. *J. Econometrics*, 79:97–127, 1997.

[14] R. B. Durand, R. A. Maller, and G. Müller. The risk return tradeoff: A COGARCH analysis in favour of Merton's hypothesis. *J. Empirical Finance.* To appear, 2009.

[15] Eberlein. Applications of generalized hyperbolic Lévy motions to finance. In O. E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick, editors, *Lévy Processes: Theory and Applications*, pages 319–336, Boston, 2001. Birkhäuser.

[16] R.F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50:987–1007, 1982.

[17] V. Fasen. Asymptotic results for sample autocovariance functions and extremes of integrated generalized Ornstein-Uhlenbeck processes. *Bernoulli*, 16(1):51–79, 2010.

[18] H. Föllmer and M. Schweizer. Hedging of contingent claims under incomplete information. In M. H. A. Davis and R. J. Elliott, editors, *Applied Stochastic Analysis*, volume 5 of *Stochastics Monographs*, pages 389–414. Gordon and Breach, London, New York, 1991.

[19] C. M. Goldie. Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.*, 1(1):126–166, 1991.

[20] C. M. Goldie and R.A. Maller. Stability of perpetuities. *Ann. Probab.*, 28:1195–1218, 2000.

[21] P.S. Hagan, D. Kumar, A.S. Lesniewski, and D.E. Woodward. Managing smile risk. *Wilmott*, Sept.:84–108, 2002.

[22] J. M. Harrison and D. M. Kreps. Martingales and arbitrage in multiperiod securities markets. *J. Econom. Theory*, 20:381–408, 1979.

[23] J. M. Harrison and S. R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stoch. Process. Appl.*, 11:215–260, 1981.

[24] S. Haug, C. Klüppelberg, A. Lindner, and M. Zapp. Method of moment estimation in the COGARCH(1,1) model. *Econom. J.*, 10:320–341, 2007.

[25] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Fin. Stud.*, 6:327–343, 1993.

[26] J. Kallsen and M.S. Taqqu. Option pricing in ARCH type models. *Math. Finance*, 8:13–26, 1998.

[27] J. Kallsen and B. Vesenmayer. COGARCH as a continuous time limit of GARCH(1,1). *Stoch. Proc. Appl.*, 119(1):74–98, 2009.

[28] H. Kesten. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131(1):207–248, 1973.

[29] C. Klüppelberg. Risk management with extreme value theory. In B. Finkenstädt and H. Rootzén, editors, *Extreme Values in Finance, Telecommunication and the Environment*, pages 101–168, Boca Raton, 2004. Chapman & Hall/CRC.

[30] C. Klüppelberg, A. Lindner, and R. Maller. A continuous time GARCH process driven by a Lvy process: stationarity and second order behaviour. *J. Appl. Prob.*, 41(3):601–622, 2004.

[31] A. Linder. Continuous time approximations to GARCH and stochastic volatility models. In T.G. Andersen, R.A. Davis, J.-P. Kreiss, and T. Mikosch, editors, *Handbook of Financial Time Series.*, pages 481–496. Springer, Berlin, 2009.

[32] A. Linder. Stationarity, mixing, distributional properties and moments of GARCH(p,q)-processes. In T.G. Andersen, R.A. Davis, J.-P. Kreiss, and T. Mikosch, editors, *Handbook of Financial Time Series.*, pages 43–70. Springer, Berlin, 2009.

[33] A. Lindner and R.A. Maller. Lévy integrals and the stationarity of generalised Ornstein-Uhlenbeck processes. *Stochastic Process. Appl.*, 115:1701–1722, 2005.

[34] D. B. Madan, P. Carr, and E.C. Chang. The variance gamma process and option pricing. *European Finance Review*, 2(1):79–105, 1998.

[35] D. B. Madan and E. Seneta. The variance gamma (VG) model for share market returns. *J. Bus.*, 63:511–524, 1990.

[36] R.A. Maller, G. Müller, and A. Szimayer. GARCH modelling in continuous time for irregularly spaced time series data. *Bernoulli*, 14:519–542, 2009.

[37] T. Mikosch. Modeling dependence and tails of financial time series. In B. Finkenstädt and H. Rootzén, editors, *Extreme Values in Finance, Telecommunication and the Environment*, pages 185–286, Boca Raton, 2004. Chapman & Hall/CRC.

[38] G. Müller. MCMC estimation of the COGARCH(1,1) model. *J. Financial Economectrics*, 8(4):481–510, 2010.

[39] G. Müller, R.B. Durand, R.A. Maller, and C. Klüppelberg. Analysis of stock market volatility by continuous-time GARCH models. In G.N. Gregoriou, editor, *Stock Market Volatility.*, pages 32–48. Chapman & Hall-CRC/Tylor and Frances, London, UK, 2009.

[40] D. B. Nelson. ARCH models as diffusion approximations. *J. Econometrics*, 45:7–38, 1990.

[41] G. Panayotov. *Three Essays on Volatility Issues in Financial Markets.* PhD thesis, University of Maryland, 2005.

[42] E. Platen and R. Sidorowicz. Empirical evidence on student-t log-returns of diversified world stock indices. *J. Statist. Theor. Pract.*, 2(2), 2008.

[43] K. Sato and M. Yamazato. Operator-selfdecomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stochastic Process. Appl.*, 17:73–100, 1984.

[44] K.-I. Sato. *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge, U.K., 1999.

[45] R. Stelzer. Multivariate COGARCH(1,1) processes. *Bernoulli*, 16(1):80–114, 2010.

[46] A. Szimayer and R.A. Maller. Finite approximation schemes for Lévy processes and their applicaiton to optimal stopping times. *Stoch. Proc. Appl.*, 117:1422–1447, 2007.

[47] S. J. Taylor. Financial returns modelled by the product of two stochastic processes: a study of daily sugar prices 1961-79. In O. D. Anderson, editor, *Time Series Analysis: Theory and Practice*, volume 1, pages 203–226. North-Holland, Amsterdam, 1982.

[48] V. Todorov and G. Tauchen. Simulation methods for Lévy driven continuous-time autoregressive moving average CARMA stochastic volatility models. *J. Bus. Econom. Statist.*, 24(4):455–469, 2006.

[49] W. Vervaat. On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Prob.*, 11:750–783, 1979.

[50] Y. Wang. Asymptotic nonequivalence of GARCH models and diffusions. *Ann. Stat.*, 30:754–783, 2002.

Centre for Mathematical Sciences, and Institute for Advanced Study, Technische Universität München, 85748 Garching, Germany

E-mail: cklu@ma.tum.de

Centre for Financial Mathematics, and School of Finance & Applied Statistics, Australian National University, Canberra, ACT

E-mail: Ross.Maller@anu.edu.au

Department of Business Administration III, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany

E-mail: szimayer@uni-bonn.de