Research Article

Farrukh Javed and Krzysztof Podgórski

Leverage Effect for Volatility with Generalized Laplace Error

Abstract: We propose a new model that accounts for the asymmetric response of volatility to positive ('good news') and negative ('bad news') shocks in economic time series – the so-called leverage effect. In the past, asymmetric powers of errors in the conditionally heteroskedastic models have been used to capture this effect. Our model is using the gamma difference representation of the generalized Laplace distributions that efficiently models the asymmetry. It has one additional natural parameter, the shape, that is used instead of power in the asymmetric power models to capture the strength of a long-lasting effect of shocks. Some fundamental properties of the model are provided including the formula for covariances and an explicit form for the conditional distribution of 'bad' and 'good' news processes given the past – the property that is important for statistical fitting of the model. Relevant features of volatility models are illustrated using S&P 500 historical data.

Keywords: Heavy Tails, Volatility Clustering, Generalized Asymmetric Laplace Distribution, Leverage Effect, Conditional Heteroskedasticity, Asymmetric Power Volatility, GARCH Models

MSC 2010: 62M10, 91G70

DOI: 10.1515/eqc-2014-0015

Received October 19, 2014; revised November 7, 2014; accepted November 26, 2014

1 Introduction

In the field of finance, it has been long observed and exhaustively documented that the data exhibit distinct features that call for more general models than linear ones based on the gaussian distribution. Among most frequently quoted non-gaussian, non-linear features are: heavy tailed distributions, clustering of volatility, asymmetry in the volatility (the leverage effect), and to a lesser degree asymmetry in the return distribution, see [8] and references therein.

In particular, the long-range dependence and volatility clustering are observed in market returns, as pointed by Mandelbrot in [18]

“...large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes.”

This stylized fact that seems to be evident in many data is sometimes associated by various authors with the empirical evidence that the absolute or squared returns are substantially more correlated at higher lags than do the returns themselves (see [5] and [8]). Moreover, the absolute and squared returns display slowly decaying autocorrelation as the lag increases.

It is also well known to financial practitioners that the vast majority of data shows systematic asymmetries. Among them two have been subject of more thorough studies, namely asymmetry in the distribution of returns, see [13] and [21] for discussions of empirical evidence, and asymmetry in the way volatility responses

to positive and negative returns. The latter, in layman's terms, means that good and bad news have different predictability for the future volatility, i.e. the volatility effect of positive response is different than that of a negative response of the same size. Such a behavior is heavily argued in finance literature under the name of the *leverage effect*, see [3] and [9].

For practical purposes such as risk management, hedging and derivative pricing, it is desirable to use a relatively simple model that can capture heavy tails and other discussed above features of financial data. One of the first volatility models for financial data extending beyond the standard linear-gaussian paradigm was proposed in the seminal paper of [7] in an attempt of modeling non-constant volatility effects (volatility clustering). The main idea is to model current, unobserved and non-constant variances of returns (volatility) through a form of autoregressive equation in which they are dependent also on past noise. Through varying volatility, a relatively simple form of non-linearity has been introduced to time series modeling. Simplicity of this model comes from two sources: firstly, conditionally on the volatility, the returns exhibit a standard linear time series structure, secondly, the volatility, when conditioned on the past, is non-random although time-varying (conditional heteroskedasticity).

If the innovation noise is gaussian, then the model conditionally on the past becomes also gaussian, see main model (2.1). This is utilized to obtain efficient estimation of the model parameters through the maximum likelihood estimation method based on the gaussian assumption about the stochastic noise that is driving randomness of the system. However, the consistency and asymptotic normality of estimation can be validated even if the errors are not gaussian in which the case it is often referred to as the quasi likelihood method, see [22] and [4].

The first rather obvious property of such models is their capability of recreating volatility clustering effects – volatility can have longer lasting effects than returns. Moreover, due to its non-linearity, volatility models potentially can account for heavier tailed returns through random and time varying volatility and the exponent (power) parameter in the volatility equation even if the errors in the leading equation are gaussian. However, as it has been observed and well documented, see [12] for a monographic account, the gaussianity assumption for errors in actual data sets is not supported by sample distribution of the residuals from the fit to the model. Namely, the residuals when scaled by empirically fit stochastic volatility possess heavier than gaussian tails. This is also illustrated in Example 1 where the S&P 500 were utilized to show that the standard A-PARCH model is not capable of accounting heavy tails in residuals.

To summarize, introducing the conditional heteroskedastic volatility allows for non-gaussian distribution of returns with heavier tails, and adding non-gaussian possibly non-symmetric distribution can also account for asymmetry in the distribution of returns, while at the same time fitting tails in the returns and residuals even better, see [8]. The class of generalized hyperbolic (GH) distributions and some of its subclasses have been utilized for this purpose in various studies, see for example, [2, 6, 17]. In order to obtain asymmetry due the leverage effect, i.e. that the 'bad' news seems to have a more prominent effect on volatility than does the 'good' news, one needs some structural changes to the model by considering a more complex, possibly non-linear, dependence on the error term. In [5], the authors address these issues by use of powers of differently weighted positive and negative parts of the error in the volatility equation. Such asymmetric power volatility models with non-gaussian errors have also been discussed in [11], see also references therein.

In this paper, we focus on a convenient GH subclass of the generalized Laplace distributions. These distributions arise as a mixture of normal distributions with stochastic gamma distributed variance and mean, lending to the name of the gamma variance/mean model. The generalized Laplace distributions have been successfully applied in finance due to their simplicity, flexibility, and good fit to empirical data (see, e.g., [14–16, 20]). The related Laplace noise can be conveniently represented as a difference of two gamma noises with different scales. We use this representation to define a new model that allows for modeling the leverage effect. In this model, the generalized autoregressive conditional heteroskedastic (GARCH) type volatility component addresses the volatility clustering phenomenon and the error distribution takes care of asymmetry and heavy tailed behavior in the data. Hence, under parsimonious representation, our proposed model will take care of all the key features observed in the financial data. We illustrate the potential of generalized Laplace distributions through analysis of S&P 500 data, spanning from January 3, 1928 to August 30, 1991, comprising a total of 17,011 values.

2 Asymmetric Power Volatility Models

The asymmetric-power autoregressive heteroskedastic volatility model was proposed in [5] as a generalization of GARCH introduced in [4]. The five parameter model is intended to account for leptokurtotic, asymmetric, and long-memory behavior of the volatility and is defined by the second equation in

$$y_t = m + \alpha y_{t-1} + \sigma \rho_t e_t, \quad \rho_t^\delta = 1 + \alpha \rho_{t-1}^\delta [(1 - \theta)^\delta e_{t-1}^{+\delta} + (1 + \theta)^\delta e_{t-1}^{-\delta}] + \beta \rho_{t-1}^\delta, \quad (2.1)$$

where $e_t \sim N(0, 1)$, while the role of the parameters is summarized in Table 1. The range of the parameters does not represent all constraints to assure stationarity of the model but proper conditions can be found in the literature, see [5]. Here we report the condition for the existence of a solution that has the first moment of ρ_t^δ finite.

Proposition 1. *A sufficient condition for existence of the strictly stationary solution to (2.1) that has finite first moment is given by the inequality*

$$(1 - \theta)^\delta + (1 + \theta)^\delta < \sqrt{\pi} \frac{1 - \beta}{\alpha} \frac{2^{1-\frac{\delta}{2}}}{\Gamma(\frac{\delta+1}{2})} \quad (2.2)$$

for the parameters $\alpha > 0$, $\beta \in [0, 1]$, and $\delta > 0$.

Parameter	Role in the model
$-\infty < m < \infty$	Location parameter for returns
$-1 < \alpha < 1$	Autoregressive parameter for returns (it should be set to zero if the random walk assumption is in place)
$\sigma > 0$	Volatility scale
$\alpha > 0$	Autoregressive parameter for volatility, the higher values the more aggressive volatility response to the market changes
$\beta > 0$	Persistence in volatility, higher values makes local volatility lasting longer
$-1 \leq \theta \leq 1$	Response to 'positive news' – asymmetry in volatility responses
$\delta > 0$	Partially controlling the tail behavior of volatility and returns as well as their dependence in time

Table 1. Role of the parameters in the model.

For the financial data to meet the efficient market assumption, one often requests that the regressive parameter α equals zero. In [5], the authors applied the model to the S&P 500 returns data and found that it provided a good fit of the data. However, there are few structural concerns regarding the proposed specification. Firstly, the asymmetry involves the positive and negative error value term that is analytically less tractable, in particular for a non-gaussian error distribution. Secondly, the use of $\delta = 1$ and $\delta = 2$ in (2.1) is naturally setting a standard deviation and variance, respectively, equation for the conditionally normal errors. However, if the data shows heavier tails which is often true for financial time series, then different powers of δ would be required to adequately describe the data, which introduces non-linearity and difficulty in interpretation. Thirdly, the model assumes that the noise process e_t is normally distributed but for real data residuals to the fitted model frequently show evident non-gaussian features as explained in the following example.

Example 1. To illustrate difficulty in capturing heavy tails, we evaluate the sample kurtosis for the returns based on S&P 500 data, spanning from January 3, 1928 to August 30, 1991 as well as the kurtosis of residuals from the fit to our volatility model. We observe in Table 2 that the residuals remain heavy tailed although the model itself removed a significant portion of mass from the tail as shown by a reduced value of the kurtosis coefficient (from 26.12 to 8.18). In contrast, the simulated data from the model that has the parameters equal to the ones fitted from the real data exhibit the kurtosis in the residuals in the vicinity of 3.0 which is the theoretical value for the normal distribution. On the other hand, we note that the gaussian error model is incapable of retrieving heavy tails of the returns (7.33 vs. 26.12). We conclude that gaussian errors are not suitable to capture heavy tails of the data and a non-gaussian noise is necessary in order to fully account for it.

Data	Kurtosis		
	Returns y_t	Residuals with volatility $\hat{\rho}_t \hat{e}_t$	Residuals $\hat{e}_t$
S&P 500 data	26.12	24.89	8.18
Simulation	7.33	7.23	2.96

Table 2. Kurtosis in the data vs. in the model.

Heavy-Tailed Asymmetric Error Distributions

A natural candidate for the distribution of errors e_t in (2.1) is the class of the generalized hyperbolic (GH) distributions which are a normal variance-mean mixture, with $\sigma > 0$ and $\mu \in \mathbb{R}$ in

$$X = \sigma\sqrt{\gamma}Z + \mu\gamma, \quad (2.3)$$

where Z is a standard normal variable independent of a non-negative mixing variable Γ having a generalized inverse gaussian (GIG) distribution so that its density is proportional to $x^{\tau-1}e^{-ax-b/x}$. In order to obtain a unique scaling in the model, one can assume that $E(\Gamma) = 1$ which leads to $E(X) = \mu$ and $\text{Var}(X) = \sigma^2$. Additionally to the four parameters, $\mu \in \mathbb{R}$, $\sigma > 0$, $\tau \in \mathbb{R}$, $a > 0$ (the final parameter of GIG, b , is obtained by setting $E(\Gamma) = 1$), one can also add a location parameter $\lambda \in \mathbb{R}$, so that the distribution of $Y = \lambda + X$ becomes fully parametrized by five parameters $\lambda, \mu, \sigma, \gamma, \tau$.

We note two important cases. Firstly, the generalized asymmetric Laplace distribution (GAL) are obtained by taking $b = 0$ and $a = \tau$, leading to the four parameters λ, μ, σ , and τ . Secondly, setting $\tau = -1/2$ leads to another important case of the normal inverse gaussian (NIG) distributions, with Γ having the inverse gamma distribution. In this case, $E(\Gamma) = 1$ yields $a = b$, leading to four parameters λ, μ, σ , and a . The importance of these subclasses lies in that these are the only two subclasses of the generalized hyperbolic distributions that are closed on convolution, see [19], allowing for arbitrary time frequency sampling.

The case of (2.1) without the autoregressive term ($a = 0$) and e_t following an NIG distribution was studied in [11]. It has been shown that the model outperforms some of the most praised models for stock returns. There are no detailed studies of the other alternative, namely, the GAL models except within a general discussion of the GH models as in [2]. However, the GAL model possesses all virtues of the NIG one, therefore it is important to demonstrate its usefulness to analyzing historical financial data, which is done next, although more extensive studies are needed and left for future research.

The model with the GAL distributed errors takes the following explicit form:

$$y_t = m + ay_{t-1} + \rho_t(\mu\gamma_t + \sigma\sqrt{\gamma_t}Z_t), \quad (2.4)$$

where γ_t are independent and identically distributed gamma distributed with the shape τ and the scale $1/\tau$. Similarly, as in the gaussian error model, it is important to determine the conditions for existence of the solution to the volatility equation. In the appendix we include a general formal argument that leads to the following result.

Proposition 2. *In the case of $\mu = 0$, a sufficient condition for existence of the strictly stationary solution ρ_t^δ in (2.1) that has finite first moment and under the model equation (2.4) is given by the inequality*

$$(1 - \theta)^\delta + (1 + \theta)^\delta < \sqrt{\pi} \frac{1 - \beta}{\alpha} \frac{2^{1-\frac{\delta}{2}}}{\Gamma(\frac{\delta+1}{2})} \frac{\tau^{\frac{\delta}{2}} \Gamma(\tau)}{\Gamma(\tau + \frac{\delta}{2})} \quad (2.5)$$

for the parameters $\alpha > 0$, $\beta \in [0, 1]$, τ , and $\delta > 0$.

We should note here that in the general asymmetric case there is no explicit formula for existence of the solution due to the fact that the moments of the positive and negative part of the error noise can be obtained only in an integral form. Consequently, numerical techniques have to be involved to practically resolve the matter of existence.

Figure 1. Autocorrelation and normal probability plot of S&P 500 stock returns.

We present some results based on analysis of the benchmark data set consisting of daily returns for the S&P 500. The autocorrelation function and the normal probability plot of the returns are reported in Figure 1. It can be seen that the raw returns deviate from normality by showing heavy tailed behavior. The autocorrelation exhibits some significant although short living serial correlation.

There are several possible strategies for fitting model (2.4) to the data, since both moments and likelihood are known for the model conditioned on the past data. For example, one can utilize a gaussian (quasi-)likelihood to fit the model, and then fit a GAL distribution to the estimated residuals. Alternatively, one can fit both the model and distribution using the GAL likelihood. Due to the presence of a Bessel function in the non-gaussian likelihood, the first procedure is computationally faster. The estimation results are summarized in Table 3, where in the first row we show fit by the gaussian likelihood and the remaining two represent the fit through the above described methods.

	m	a	α	β	θ	δ	σ	μ	τ
Gaussian	0.0003	0.14	0.09	0.91	0.33	1.73	0.014	–	–
$\text{GAL}_{q\text{-mle}}$	–	–	–	–	–	–	0.009	–0.0008	2.21
GAL_{mle}	0.0004	0.12	0.05	0.91	0.37	1.52	0.098	–	2.22

Table 3. Estimated parameters of the asymmetric power autoregressive conditionally heteroskedastic model.

In Figure 2, one can observe that the choice of the GAL likelihood helped to account for the tail and asymmetry in the data. There are very few observations which are off at the tails in these probability plots but their presence could be explained by some outliers in the historical data.

3 New Asymmetric Volatility Model with GAL Errors

Positive news affects volatility differently from negative one. In the asymmetric power model, the news on any given day is summarized either as positive or as negative by taking the absolute value of the stochastic innovation term as shown in the second equation of (2.1). However, in reality both negative and positive news are reported daily. The important property of generalized asymmetric Laplace distributions is that they offer a way to model both negative and positive news processes and not only their daily overall balance. It is due to the fact that they can be represented not only as the mean variance normal mixture (2.3) but also as

Figure 2. Sample distribution of residuals. (a) qq-plot against the gaussian distribution, (b) qq-plot against the fitted GAL, (c) histogram of the residuals against the fitted GAL density.

a difference of two gamma distributed variables

$$e_t = \sigma \left(\kappa \gamma_t^+ - \frac{\gamma_t^-}{\kappa} \right), \quad (3.1)$$

where γ_t^+ and γ_t^- are two independent and identically distributed gamma random variables with the same shape parameter τ and scale $1/\tau$ so that their expected values are equal to one. According to this at any given time point t the innovation e_t is a mixture of ‘good’ (γ_t^+) and ‘bad’ (γ_t^-) news. The non-linear conditional volatility model is adopted from (2.1) by taking gamma processes γ_t^+ and γ_t^- , instead of, as previously, overall balance of news, i.e. positive and negative parts of random noise e_t . Thus the expression for ρ_t^δ becomes

$$\rho_t^\delta = 1 + \alpha \rho_{t-1}^\delta ((1 - \theta)^\delta \gamma_{t-1}^{+\delta} + (1 + \theta)^\delta \gamma_{t-1}^{-\delta} + \beta). \quad (3.2)$$

We observe that both γ_{t-1}^+ and γ_{t-1}^- can be non-zero and we also avoid using the absolute value which can be analytically troublesome, specially for asymmetric distributions of e_t . While we have included the power parameter δ in the model, there is already another parameter that can fit the tail behavior, namely, shape τ of the gamma distribution. Therefore the models with a natural specified choice of $\delta = 1$ or $\delta = 2$ are of the principal interest. By using the shot-process representation of gamma distribution it is also possible to

model ρ_t through a different value of the shape parameter than the one used for y_t – an approach that is not discussed here.

A natural question is for what values of the parameters there exists a jointly stationary solution ρ_t and y_t . Here is a result that provides such a condition. Its proof is discussed in the appendix.

Proposition 3. *A sufficient condition for existence of the strictly stationary solution to (3.2) that has finite first moment is given by the inequality*

$$(1 - \theta)^\delta + (1 + \theta)^\delta < \frac{1 - \beta}{\alpha} \frac{\tau^\delta \Gamma(\tau)}{\Gamma(\tau + \delta)}$$

for the parameters $\tau > 0$, $\alpha > 0$, $\beta \in [0, 1]$, $\theta \in [0, 1]$, and $\delta > 0$.

Corollary 1. *In the proposed model, the tails can be controlled by shape parameter τ , thus it is natural to consider:*

(i) $\delta = 1$ in which the case the sufficient condition is independent of τ as it becomes

$$2\alpha < 1 - \beta,$$

(ii) $\delta = 2$ with the stationarity condition given by

$$2\left(1 + \frac{1}{\tau}\right)(1 + \theta^2)\alpha < (1 - \beta).$$

Conditional Distributions on the Past and Estimation Strategies

For any new model of practical importance one has to address model fitting. Presentation of a complete approach to this complex problem extends beyond the purpose of this note. Here we present a key property that can be utilized when fully addressing this topic. We start with reviewing fundamentals of the likelihood based approach to estimation for the conditional heteroskedastic volatility models.

Consider a sequence of past data $\mathbf{y}_t = (y_s : s \leq t)$ and corresponding error terms $\mathbf{e}_t = (e_s : s \leq t)$ that are independent and identically distributed with a density h_θ . Then the likelihood for the observed $y_1, \dots, y_t$ can be approximately written as

$$L(\theta; y_1, \dots, y_t | \mathbf{y}_0) = f_\theta(y_1, \dots, y_t | \mathbf{y}_0) = \prod_{i=1}^t \int h_\theta\left(\frac{y_i - m - ay_{i-1}}{\rho}\right) g_\theta(\rho | \mathbf{e}_{i-1}) \rho^{-1} d\rho, \quad (3.3)$$

where we assume that the past $\mathbf{y}_t$ is uniquely defined by the past error terms $\mathbf{e}_t$. Here $g_\theta(\rho | \mathbf{e})$ is the conditional distribution of ρ_t given the past $\mathbf{e}_t = \mathbf{e}$.

For the classical asymmetric power models (both with gaussian and non-gaussian noise), the ρ_t depends deterministically on the past (this is the reason for coining the term ‘conditional heteroskedasticity’). As the result the above likelihood simplifies to

$$L(\theta; y_1, \dots, y_t | \mathbf{y}_0) = \prod_{i=1}^t h_\theta\left(\frac{y_i - m - ay_{i-1}}{\rho_i}\right) \frac{1}{\rho_i},$$

where ρ_t is the value computable from the past values of errors $\mathbf{e}_{t-1}$. In the newly proposed model, the volatility ρ_t remains random when it is conditioned on the past history of returns. Therefore in order to obtain estimation methods based on the likelihood, the conditional distribution $g_\theta(\rho | \mathbf{e})$ has to be obtained.

More specifically, it is important to express distribution of the ‘bad’ and ‘good’ components of the volatility ρ_t conditionally on the past data. It can be argued that the problem reduces to obtaining the conditional distribution of Γ_1 and Γ_2 given that $E = e$ is observed in $E = \kappa\Gamma_1 - \Gamma_2/\kappa$, where Γ_1 and Γ_2 are independent gamma distributed random variables with scale 1 and shape τ . These distributions are easiest expressed by a three parameter family of distributions that we have termed the *tilted gamma distributions* and which are characterized by the densities proportional to $x^{\tau-1}(x+a)^\alpha e^{-x}$, $x > 0$. Our main result about the conditional distribution concludes this work, while its standard proof is given in Appendix B.

Proposition 4. *If $e < 0$, then the distribution of Γ_1 given that $E = e$ is that of $X/(1 + \kappa^2)$, where X has the tilted gamma distribution with shapes τ , $\alpha = \tau - 1$, and tilting $a = (\kappa + \kappa^{-1})|e|$. When on the other hand $e \geq 0$, the distribution of Γ_1 given that $E = e$ is that of $(X + (\kappa + \kappa^{-1})e)/(1 + \kappa^2)$.*

The distribution of Γ_2 with parameter κ given that $E = e$ is equal to the conditional distribution of Γ_1 given that $E = -e$ with parameter κ replaced by κ^{-1} .

Despite having the explicit form of the density $g_\theta(\rho | \mathbf{e})$, effectively evaluating the likelihood (3.3) could be still a challenging task due to the necessity of evaluating the integral with respect to the conditional density. In a simplified approach one could replace this integral by considering the conditional expected values $\bar{\rho}_t = E(\rho_t | \mathbf{e}_t)$ and maximizing the following simplified ‘likelihood’:

$$L(\boldsymbol{\theta}; y_1, \dots, y_t | \mathbf{y}_0) = \prod_{i=1}^t h_\theta \left(\frac{y_i - m - ay_{i-1}}{\bar{\rho}_i} \right) \frac{1}{\bar{\rho}_i}.$$

For this approach, one can utilize the following fact that follows from standard integration using tilted gamma densities.

Corollary 2. *The first conditional moment of Γ_1 given that $E = e$ can be evaluated according to*

$$E(\Gamma_1 | E = e) = \frac{\sqrt{\pi}}{(\kappa^2 + 1)K_{\tau-\frac{1}{2}}((\kappa + \kappa^{-1})\frac{|e|}{2})} \begin{cases} W_{\frac{-1}{2}, \tau}((\kappa + \kappa^{-1})|e|)\tau, & e < 0, \\ W_{\frac{1}{2}, \tau}((\kappa + \kappa^{-1})e), & e > 0. \end{cases} \quad (3.4)$$

Here, we use the Whittaker function that can be expressed for $\mu, \kappa \in \mathbb{R}$, $\kappa < \mu + \frac{1}{2}$, $x > 0$:

$$W_{\kappa, \mu}(x) = \frac{x^{\frac{1}{2}-\mu} e^{-\frac{x}{2}}}{\Gamma(\mu - \kappa + \frac{1}{2})} \int_0^\infty e^{-u} u^{\mu-\kappa-\frac{1}{2}} (u+x)^{\mu-\frac{1}{2}+\kappa} du,$$

see [10] and [1].

Conclusions

In the paper, volatility models are considered that account for asymmetry, heavy tail, and different effects of ‘good’ and ‘bad’ news on economic data. Using S&P 500 returns data, it is demonstrated that the asymmetric power autoregressive conditionally heteroskedastic volatility model has difficulty in addressing distributional effects unless one considers non-gaussian errors. It is also argued that the generalized asymmetric Laplace errors are as attractive in modeling financial data as the normal inverse gaussian errors that were presented in [11].

A new attractive volatility model is introduced that also features generalized asymmetric Laplace errors. It is unique to this class since it utilizes the representation of errors as a difference of scaled gamma variables. This representation allows for a natural model for the ‘good’ and ‘bad’ news processes. The fundamental fact about the distributions of these two processes given the past is presented that can be utilized for effective model fitting to data. Derivation of data fitting methods and further investigation of the model is left for future studies.

A Existence of Stationary Solutions

In this appendix, we provide with the details of the argument for existence of stationary solutions under general heteroskedastic autoregressive models for volatility.

Our results presented in the paper are consequences of the following lemma.

Lemma A.1. Let λ_k be independent identically distributed non-negative random variables with finite first moment $m = E\lambda_k$ smaller than 1. Then the time series

$$\rho_t^\delta = \alpha_0 \left(1 + \sum_{k=1}^{\infty} \lambda_{t-1} \cdots \lambda_{t-k} \right) \quad (\text{A.1})$$

is well-defined in the mean sense and satisfies the following recurrent equation:

$$\rho_t^\delta = \alpha_0 + \rho_{t-1}^\delta \lambda_{t-1}, \quad t = \dots, -1, 0, 1, \dots \quad (\text{A.2})$$

Proof. The proof is a simple observation of the fact that the norm of the series is a geometric series with the quotient m :

$$E \sum_{k=1}^{\infty} \lambda_{t-1} \cdots \lambda_{t-k} = \sum_{k=1}^{\infty} m^k = \frac{m}{1-m},$$

as long as m is smaller than 1. □

In the paper we considered three cases of the above model for ρ_t . Namely, in Proposition 1 the well-known case is considered for which

$$\lambda_t = (1 - \theta)^\delta e_{t-1}^{+\delta} + (1 + \theta)^\delta e_{t-1}^{-\delta},$$

in Proposition 2 the asymmetric power model with symmetric generalized Laplace errors yields

$$\lambda_t = \gamma_t^{\frac{\delta}{2}} ((1 - \theta)^\delta e_{t-1}^{+\delta} + (1 + \theta)^\delta e_{t-1}^{-\delta}),$$

and, finally, the new model with gamma modeled ‘bad’ and ‘good’ news dealt in Proposition 3

$$\lambda_t = (1 - \theta)^\delta \gamma_{t-1}^{+\delta} + (1 + \theta)^\delta \gamma_{t-1}^{-\delta} + \beta.$$

Evaluation of the expectations based on the following well-known relation for the moments of a gamma variable,

$$E\gamma^\delta = \frac{\Gamma(\tau + \delta)}{\tau^\delta \Gamma(\tau)}, \quad (\text{A.3})$$

proves the conditions stated in the propositions for each of the above cases.

B Conditional Distributions of ‘Good’ and ‘Bad’ News

We collect here some basic properties of the conditional distributions of ‘good’ and ‘bad’ news given overall value of error term. The error E can be written as

$$E = \kappa \Gamma_1 - \frac{\Gamma_2}{\kappa}, \quad (\text{B.1})$$

where Γ_1 and Γ_2 are independent gamma distributed random variables with scale 1 and shape τ . In this section we study the conditional distribution of Γ_1 and Γ_2 given that $E = e$ is observed.

Let us consider first the distribution of Γ_1 given that $E = e$. By the change of variables

$$\begin{aligned} \gamma &= \gamma_1, \\ e &= \kappa \gamma_1 - \frac{\gamma_2}{\kappa}, \end{aligned}$$

mapping $(\gamma_1, \gamma_2) \in [0, \infty)^2$ to $(\gamma, e) \in [0, \infty) \times \mathbb{R}$, we obtain the joint density of Γ_1 and E :

$$\begin{aligned} f_{\Gamma_1, E}(\gamma, e) &= \kappa f_{\Gamma_1}(\gamma) f_{\Gamma_2}(-\kappa\gamma + \kappa^2\gamma) \\ &= \frac{\kappa^{2\tau-1}}{\Gamma^2(\tau)} \left(\gamma \left(\gamma - \frac{e}{\kappa} \right) \right)^{\tau-1} \exp(-(1 + \kappa^2)\gamma + \kappa e), \quad \gamma > 0, e < \kappa\gamma. \end{aligned}$$

Here and in what follows functions are assumed to vanish outside the specified range of their arguments, and if the range is not specified, then it is assumed to be the entire real line.

Since the density of a generalized asymmetric Laplace distribution is given by

$$f_E(e) = \frac{\exp((\kappa - \kappa^{-1})\frac{e}{2})}{\pi^{\frac{1}{2}}\Gamma(\tau)} \left(\frac{|e|}{\kappa + \kappa^{-1}}\right)^{\tau - \frac{1}{2}} K_{\tau - \frac{1}{2}}\left((\kappa + \kappa^{-1})\frac{e}{2}\right),$$

the conditional density of $\Gamma_1 | E = e$ can then be written as

$$f_{\Gamma_1 | E}(\gamma | e) = C_e \cdot \left(\gamma \left(\gamma - \frac{e}{\kappa}\right)\right)^{\tau - 1} \exp(-(1 + \kappa^2)\gamma), \quad \gamma > 0 \vee \frac{e}{\kappa},$$

where

$$C_e = \frac{\pi^{\frac{1}{2}}(\kappa^3 + \kappa)^{\tau - \frac{1}{2}}}{\Gamma(\tau)} \frac{e^{(\kappa + \kappa^{-1})\frac{e}{2}}}{|e|^{\tau - \frac{1}{2}} K_{\tau - \frac{1}{2}}((\kappa + \kappa^{-1})\frac{e}{2})}.$$

We recognize that this is the tilted gamma distribution with the tilting parameter $e(\kappa + 1/\kappa)$.

Funding: The authors were supported by the Riksbankens Jubileumsfond Grant Dnr: P13-1024:1 and the Swedish Research Council Grant Dnr: 2013-5180.

References

- [1] Digital library of mathematical functions, 2010, <http://dlmf.nist.gov>.
- [2] O. Barndorff-Nielsen and N. Shephard, Non-gaussian ornstein-uhlenbeck-based models and some of their uses in financial economics (with discussion), *Journal of the Royal Statistical Society B* **63** (2001), 167–241.
- [3] F. Black, Studies in stock price volatility changes, in: *Proceedings of the 1976 Business Meeting of the Business and Economics Section*, American Statistical Association, (1976), 177–181.
- [4] T. Bollerslev, Generalized conditional autoregressive heteroskedasticity, *Journal of Econometrics* **31** (1986), 307–327.
- [5] Z. Ding, C. Granger and R. Engle, A long memory property of stock market returns and a new model, *Journal of Empirical Finance* **1** (1993), 83–106.
- [6] E. Eberlein and U. Keller, Hyperbolic distributions in finance, *Bernoulli* **1** (1995), 281–299.
- [7] R. Engle, Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **45** (1982), 987–1007.
- [8] J. Eric, S. Poon and M. Rockinger, *Financial Modelling Under Non-Gaussian Distributions*, Springer, London, 2007.
- [9] B. Feunou and R. Tédongap, A stochastic volatility model with conditional skewness, *Journal of Business & Economic Statistics* **30** (2012), no. 4, 576–591.
- [10] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, Orlando, 1980.
- [11] M. B. Jensen and A. Lunde, The NIG-S&ARCH model: A fat-tailed, stochastic, and autoregressive conditional heteroskedastic volatility model, *Econometrics Journal* **4** (2001), 167–342.
- [12] E. Jondeau, S.-H. Poon and M. Rockinger, *Financial Modeling Under Non-Gaussian Distributions*, Springer, London, 2007.
- [13] E. Jondeau and M. Rockinger, Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements, *Journal of Economic Dynamics and Control* **27** (2003), no. 10, 1699 – 1737.
- [14] S. Kotz, T. Kozubowski and K. Podgórski, *The Laplace Distribution and Generalizations. A Revisit with Applications to Communications, Economics, Engineering and Finance*, Birkhäuser, Basel, 2001.
- [15] T. Kozubowski and K. Podgórski, Laplace probability distributions and related stochastic processes, in: *Probability. Interpretation, Theory and Applications*, Nova Science Publishers, New York (2012), 105–145.
- [16] T. J. Kozubowski and K. Podgórski, Asymmetric laplace laws and modeling financial data, *Mathematical and Computer Modelling* **34** (2001), 1003–1021.
- [17] D. B. Madan, P. Carr and E. C. Chang, The variance gamma process and option pricing, *Review of Finance* **2** (1998), 74–105.
- [18] B. B. Mandelbrot, The variation of certain speculative prices, *Journal of Business* **36** (1963), 392–417.
- [19] K. Podgórski and J. Walin, Convolution invariant subclasses of generalized hyperbolic distributions, technical report, Lund University, 2013.
- [20] E. Seneta, Fitting the variance-gamma model to the financial data, *Journal of Applied Probability* **41** (2004), 177–187.
- [21] P. Verhoeven and M. McAleer, Fat tails and asymmetry in financial volatility models, *Mathematics and Computers in Simulation* **64** (2004), no. 3–4, 351–361.
- [22] A. A. Weiss, Asymptotic theory for arch models: Estimation and testing, *Econometric Theory* **2** (1986), no. 1, 107–131.