Modeling heteroscedastic, skewed and leptokurtic returns in discrete time

Ivivi Joseph Mwaniki

Abstract

Popular models of finance, fall short of accounting for most empirically found stylized features of financial time series data, such as volatility clustering, skewness and leptokurtic nature of log returns. In this study we propose a general framework for modeling asset returns which account for serial dependencies in higher moments and leptokurtic nature of scaled GARCH filtered residuals. Such residuals are calibrated to normal inverse Gaussian and hyperbolic distribution. Dynamics of risky assets assumed in Black Scholes model, Duans GARCH model and other benchmark models for contract valuation, are shown to be nested in the proposed framework.

Keywords: Stylized facts; Normal inverse Gaussian; GARCH model; Hyperbolic distribution; leptokurtic returns

1 Introduction

Uncertainty is central to much of modern finance theory. In option pricing for example, the uncertainty associated with future return of the underlying asset, is
the most important determinant in the pricing function. Popular models such as
the Black and Scholes [2] model, based on the geometric Brownian motion have very
nice mathematical properties which have been extensively used to value contracts.
Empirical evidence suggest that the dynamics of the underlying process under the
physical measure $\mathbb{P}$ follow a more complicated process than the standard geometric
Brownian motion with constant volatility.

Various studies have shown that the normal distribution does not accurately de-
scribe observed stock return data. In literature, for example [8], [1], [5] and references
therein, it is proposed that daily log returns, could be modeled by an exponential
Lévy processes and geometric Lévy process, generating a lot of literature applied
in pricing derivatives. Moreover, other studies suggest that a more improved model
would include both stochastic volatility and a jump component (see [4]). Over the
past several decades, several stylized facts have emerged about the statistical behav-
ior of speculative market daily returns such as aggregational Gaussianity, volatility
clustering, changing variance, Taylor effect, leptokurtic residuals etc, see [17], [6],
[18] and [16] for further documentation.

A typical finding concerning the return characteristic is that one period asset
return, conditional on the most up-to-date information, exhibit a fat tail behaviour
in addition to varying second order moments. The ARCH family of models intro-
duced by [9] and generalized by [3] has in the recent years gained prominence
for modeling such dynamics. In the last few years, much interest has been given
to the discrete-time GARCH option pricing models for instance see [15], [10], and
[1]. Option pricing in GARCH models has been typically done using the local risk
neutral valuation relationship (LRNVR) pioneered by [7]. The crucial assumptions
in his GARCH model construction, are the conditional normal distribution of the
asset returns under the physical probability measure $\mathbb{P}$ and the invariance of the
conditional volatility to the change of measure. The main objective of this study
is to propose a general framework for modeling the underlying uncertainty driving
eheteroscedastic and leptokurtic daily stock market returns.

This article is organized as follows, the next section presents ACH type model and
class of generalized hyperbolic distribution for modeling some of the basic stylized
facts of returns. In section three, we state the proposed general framework followed
by several examples of popular benchmark models nested within the framework.
Empirical analysis and parameter estimation are stated in section four, and section
five concludes the study.
2 Preliminary considerations

Stock returns tend to exhibit a significant serial dependency in second moments. ARCH class models and several class of statistical distributions say NIG etc are known to models skewness and excess kurtosis of log returns.

2.1 ARCH models: Changing variance

ARCH model models have become popular for modeling financial time series because they are able to account for several empirical features like volatility clustering and leptokurtic in the distribution of returns. Most ARCH-type models involve a sequence of innovations whose variance is random. Conditioned on the past the variance depends only on the previous innovations and the previous conditional variances, and standard Wiener process generate the filtration. Most studies of daily stock returns using GARCH models and conditioned on normal distribution for the innovations; the re-scaled residuals showed excess kurtosis which violates the normality assumption. We investigate AR(1)-GARCH(1,1) model conditioned on normal distribution. Thus

\[
X_t = \log e S_t - \log e S_{t-1} = \hat{\mu} + \phi X_{t-1} + \sigma_t (Z_t + \xi_t), Z_t \sim N(0, 1)
\]

Parameter estimates \(\hat{\mu}, \phi, \omega, a, b\) are presented in Table 1 and Table 2.

A typical finding concerning the return characteristic is that one period asset return conditional on the most up to date information, continues to exhibit a fat tailed behaviour. This fat tailed behaviour is also known as conditionally leptokurtic, and can be modeled by a limiting class of generalized hyperbolic distribution.

2.2 Generalized hyperbolic distribution

The probability density function of the one-dimensional Generalized Hyperbolic distribution is given by the following:

\[
f_{GH}(x; \alpha, \beta, \delta, \mu, \lambda) = \frac{(\gamma/\delta)^{\lambda}}{\sqrt{2\pi K_\lambda(\delta\gamma)}} \cdot \frac{K_{\lambda-\frac{1}{2}}(\alpha \sqrt{\delta^2 + (x - \mu)^2})}{(\sqrt{\delta^2 + (x - \mu)^2}/\alpha)^{\frac{\lambda}{2} - \lambda}} \cdot e^{\beta(x - \mu)}
\]
where $K_\lambda$ is a modified Bessel function of the third kind with the index $\lambda$.

$$K_\lambda(\omega) = \frac{1}{2} \int_0^\infty \exp \left[ -\frac{\omega}{2} (v^{-1} + v) \right] v^{\lambda-1} dv \quad (2.3)$$

Many distributions are obtained as limiting distributions of the generalized hyperbolic distribution and by varying parameter $\lambda$ to obtain subclasses for example, hyperbolic distribution and normal inverse Gaussian.

### 2.2.1 Hyperbolic distributions

When $\lambda = 1$, we obtain the subclass of hyperbolic distribution with probability density function (see [8],)

$$f_{\text{hyp}}(x; \alpha, \beta, \delta, \mu) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha\delta K_1 \left( \frac{\delta}{\sqrt{\alpha^2 - \beta^2}} \right)} \exp \left[ -\alpha \sqrt{\delta^2 + (x - \mu)^2} + \beta (x - \mu) \right]. \quad (2.4)$$

The mean and variance of hyperbolic function given respectively by the followings

$$E(X) = \mu + \frac{\beta\delta}{\sqrt{\alpha^2 - \beta^2}} K_2(\eta)$$

$$Var(X) = \delta^2 \left( \frac{K_2(\eta)}{\eta K_1(\eta)} + \frac{\beta^2}{\alpha^2 - \beta^2} \left[ \frac{K_3(\eta)}{K_1(\eta)} - \left( \frac{K_2(\eta)}{K_1(\eta)} \right)^2 \right] \right)$$

where $\eta = \delta \sqrt{\alpha^2 - \beta^2}$. The first two parameters $\alpha$ and $\beta$ determine the shape of the distribution, while the other two $\delta$ and $\mu$ are scale and location parameters.

### 2.2.2 NIG distribution

A random variable $X \sim NIG(\alpha, \beta, \delta, \mu)$ if (see [1])

$$f_{\text{NIG}}(x; \alpha, \beta, \delta, \mu) = \frac{\alpha}{\pi} \exp \left( \delta \left[ \sqrt{\alpha^2 - \beta^2} + \beta \zeta(x) \right] \right) \frac{K_1(\alpha \delta \sqrt{1 + \zeta(x)^2})}{\sqrt{1 + \zeta(x)^2}} \quad (2.7)$$

where $\zeta(x) = (x - \mu)/\delta$ and $K_1$ is the modified Bessel function of third kind, with the index 1.

$$K_1(\omega) = \frac{1}{2} \int_0^\infty \exp \left[ -\frac{\omega}{2} (v^{-1} + v) \right] dv$$
Log likelihood function for MLE of parameters

\[
\mathcal{L}_{NIG}(x|\alpha, \beta, \delta, \mu) = -n \ln \left( \frac{\pi}{\mu} \right) + n(\delta \sqrt{\alpha^2 - \beta^2}) - \frac{1}{2} \sum_{i=1}^{n} \log(1 + \zeta(x_i)^2) \\
+ \beta \delta \sum_{i=1}^{n} \zeta(x_i) + \sum_{i=1}^{n} \log(K_1(\delta \alpha \sqrt{(1 + \zeta(x_i)^2)})
\]

(2.8)

3 Modeling underlying process

Most of the notation and model framework in this paper, is a slight modification of the skewed and leptokurtic generalized GARCH framework proposed by [18] and references therein. In that study, the dynamics of log returns \( X_t \) are specified as

\[
X_t = m_t(\cdot; \theta_m) + \sigma_t \varepsilon_t \\
\sigma_t^2 = g(\sigma_s^2; \varepsilon_s; -\infty < s \leq t - 1; \theta_h) \\
\varepsilon|\mathcal{F}_{t-1} \sim D(0, 1; \theta_D)
\]

where \( \mathcal{F}_{t-1} \) is the information set containing all information up to and including time \( t - 1 \), \( m_t(\cdot; \theta_m) \) denote the conditional mean, governed by a set of parameters \( \theta_m \), \( D(0, 1; \theta_D) \) denote a zero mean and a unit variance distribution function, allowed to depend on a set of parameters \( \theta_D \).

Let \( (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{P}) \) be a stochastic basis describing the uncertainty of the economy. We refer to \( \mathbb{P} \) as the physical probability measure and \( \mathcal{F}_t \) represent the information flow driven by Wiener process \( W = (W_t)_{t \in [0,T]} \) and Lévy proces \( L = (\mathcal{L}_t)_{t \in [0,T]} \). Let \( S_t \) be the price of a stock at time \( t \) adapted to the natural filtration \( \mathcal{F}_t \). Define daily log return as \( X_t = \log \left( \frac{S_t + d_t}{S_{t-1}} \right) \), \( t = 1, 2, \ldots \) where \( d_t \) denote dividends at time \( t \) and \( X_t \) the continuously

**Theorem 3.1.** We propose the following model for asset returns under \( \mathbb{P} \). Let

\[
X_t = \log \left( \frac{S_t + d_t}{S_{t-1}} | \mathcal{F}_{t-1} \right)
\]
where \( d_t \) is one period dividends paid, then

\[
X_t = m_t(\theta_m) + \sigma_t(W_t + \xi_t), \quad (3.1)
\]

\[
\sigma_t^2 = g(\sigma_s^2, \vartheta_s; -\infty < s \leq t - 1; \theta_{\sigma}),
\]

\[
\vartheta_t \sim \text{i.i.d.}(0, 1), \quad \xi_t \sim \text{i.i.d.}(0, \theta_{\xi})
\]

where \( m_t(\theta_m) \) denote the mean function, \( \sigma_t(\theta_{\sigma}) \) denote the variance process and \( \theta = (\theta_m, \theta_{\sigma}, \theta_{\xi}) \).

### 3.1 Examples from Benchmark models

Without loss of generality, for all the subsequent examples we assume \( d_t = 0 \) \( \forall t > 0 \)

**Example 3.1. Geometric Brownian motion**

Let \( dS_t = \mu S_t dt + \sigma S_t dW_t \) be the stochastic differential equation modeling the uncertainty of the underlying process, then

\[
S_t = S_0 \exp\left(\left[\mu - \frac{\sigma^2}{2}\right] t + \sigma \sqrt{t} Z\right), \quad Z \sim N(0, 1)
\]

\[
\Rightarrow X_t = \ln(S_t/S_{t-1}) = \left[\mu - \frac{\sigma^2}{2}\right] + \sigma Z
\]

thus

\[
X_t = m_t(\theta_m) + \sigma_t(Z + \varpi), \quad m_t = \left[\mu - \frac{\sigma^2}{2}\right], \quad \sigma_t = \sigma, \quad Z \sim \text{i.i.d.} N(0, 1)
\]

**Example 3.2. Jump-diffusion model**

Merton(1976) introduced and analyzed one of the first models with both jump and diffusion term for pricing of derivative securities. Merton jump-diffusion model can be specified through the SDE.

\[
\frac{dS_t}{S_t} = \mu dt + \sigma dW_t + (J_t - 1) dN_t
\]

where the jump sizes are \( J_t \) are identically distributed and mutually independent. He also assumed that the three processes \((W_t)_{t \geq 0}, (N_t)_{t \geq 0} \) and \((J_t)_{t \geq 0} \) are independent. Let \( \sum_{j=1}^{N(t)} (Y_j - 1) \) where \( Y_1, Y_2, \ldots \) are random variables and \( N(t) \) is a counting process. The following expression solves the SDE.
\[
S(t) = S(0) \exp \left[ \left( \mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right] \prod_{j=1}^{N(t)} Y_j
\]

which implies

\[
\log_e \left( \frac{S(t)}{S(0)} \right) = \left( \mu - \frac{\sigma^2}{2} \right) t + \sigma W_t + \sum_{j=1}^{N_t} \log_e Y_j
\]

\[
= \left( \mu - \frac{\sigma^2}{2} \right) t + \sigma \left( W_t + \frac{\varpi}{\sigma} \sum_{j=1}^{N_t} \log_e Y_j \right)
\]

Given any date \( t \geq 0 \) and a holding period of length \( h > 0 \), the returns \( X_t(h) \) over the period \( [t, t+h] \) is a model given by

\[
X_t(h) = \begin{cases} 
x, & \text{if } K = 0; 
x + y_1 + ... + y_k, & \text{if } K \geq 1. 
\end{cases} \tag{3.3}
\]

where \( x \sim N(\alpha h, \sigma^2 h) \), \( \alpha = (\mu - (\sigma^2/2)) \), \( y_1, ..., y_k \) is an i.i.d. sequence with common distribution say \( G \) and \( K \) is Poisson with parameter \( \lambda h \), \( \lambda > 0 \) for \( k = 0, 1, 2, ... \), we have \( \text{Prob}(K = k) = \exp(-\lambda h)(\lambda h)^k/k \). If \( Y_j \) have the log normal distribution \( LN(a, b^2) \) then \( \log_e Y_j \sim N(a, b^2) \) and \( \sum \log Y_j \sim N(an, b^2n) \) which is the same as \( \sum \log Y_j \sim an + b\sqrt{nN}(0, 1) \). Thus

\[
X_t = \log_e \left( \frac{S_t}{S_{t-1}} \right)
\]

\[
= \left( \mu - \frac{\sigma^2}{2} \right) + \sigma \left( Z_t + \frac{\eta}{\sigma} \sum_{j=1}^{K} \log_e Y_j \right)
\]

\[
= m(\theta_m) + \sigma t (Z_t + \varpi \xi) \text{ where } \xi = \sum_{j=1}^{K} \log_e Y_j \tag{3.4}
\]

**Example 3.3.** Duan(1995) GARCH model
Discrete time economy, one period rate of return assumed to be conditionally log-normally distributed under \( \mathbb{P} \). Let \( r \) be constant one period risk free rate of return,
λ be constant unit risk premium then

\[ X_t = \log(S_t/S_{t-1}) = r + \lambda \sigma_t - \frac{\sigma_t^2}{2} + \sigma_t^2 Z, \text{ where} \]

\[ \sigma_t^2 = \omega + \sum_{j=1}^{q} \alpha_j \sigma_{t-j}^2 Z_{t-j} + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \quad Z_{t-j} \sim i.i.d. N(0,1) \forall j \]

\[ = m_t(:, \theta_m) + \sigma_t(\vartheta_t + \varpi \xi) \text{ where } \sigma_t \sim GARCH(p,q) \quad (3.5) \]

**Example 3.4.** It is well known from empirical studies that \( X_t \) can be represented as \( X_t = \mu_t + \varepsilon_t + \xi_t \) where \( \mu_t \) is a mean function and \( \varepsilon_t, \xi_t \) are the two components of the error term (see for instance [13],[11],[12],[14]). Moreover, define a \( p^\text{th} \) order autoregressive process \( \{X_t, t \geq 0\} \) with GARCH(\( p,q \)) error as

\[ X_t = \mu_t + \varepsilon_t + \xi_t, \text{ where} \]

\[ \mu_t = \sum_{r=1}^{p} \phi_r X_{t-r} + \mu, t \in \mathbb{Z}^+ \]

\[ \varepsilon_t + \xi_t = \sigma_t(Z_t + \sigma L_t), \quad Z_t, \quad \text{and } L_t \sim i.i.d(0,1), \quad Z_0 = 0, \quad L_0 = 0 \]

\[ \sigma_t^2 = \text{GARCH}(p,q), p,q \in \mathbb{Z}^+ \]

where \( Z_t \) and \( L_t \) are identically and independently distributed random variables. A general time series model for log returns would be

\[ X_t = \mu_t + \sigma_t(Z_t + \sigma L_t), \quad Z_t \sim N(0,1), \quad L_t \in GH \quad (3.6) \]

\[ = m(:, \theta_m) + \sigma_t(Z_t + \varpi L_t) \]

**4 Empirical Analysis**

We investigate the statistical properties under the objective measure \( \mathbb{P} \) model’s ability to explain observed market share prices.
4.1 Data description and parameter estimates

We apply our framework to stock indices sample from two top GDP countries, i.e. S&P500 from New York Stock exchange and CAC40 of Paris stock exchange. The financial time series data consist of S&P500 and CAC40 index daily closing adjusted price from January 2, 2001 through December 31, 2014. Daily adjusted closing prices were used to determine daily log returns $X_t$, $t = 0, 1, 2, ..., n$. Let $S_j$ be the price on day $j$, $j = 0, 1, 2, ..., n - 1$. Sample increments of log returns is defined by $X_j = \log S_j - \log S_{j-1}$, $j = 1, 2, ..., n - 1$. GARCH models are well known to be the best performing models to describe evolution of volatility, a satisfactory statistical fit is provided when the distribution of the filtered historical residuals is non-Gaussian. We fit $AR(1) + GARCH(1, 1)$ model conditioned on normal distribution. All models parameters are estimated by numerical maximum likelihood routine and the significant parameters are reported in Table 1 and Table 2.

<table>
<thead>
<tr>
<th>Table 1: $AR(1) - GARCH(1, 1)$ model parameter estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S&amp;P500</strong></td>
</tr>
<tr>
<td>$\bar{\mu}$</td>
</tr>
<tr>
<td>$\phi$</td>
</tr>
<tr>
<td>$\omega$</td>
</tr>
<tr>
<td>$a$</td>
</tr>
<tr>
<td>$b$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: $AR(1) - GARCH(1, 1)$ model parameter estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CAC40</strong></td>
</tr>
<tr>
<td>$\bar{\mu}$</td>
</tr>
<tr>
<td>$\phi$</td>
</tr>
<tr>
<td>$\omega$</td>
</tr>
<tr>
<td>$a$</td>
</tr>
<tr>
<td>$b$</td>
</tr>
</tbody>
</table>

The standardized filtered residuals are known to be uncorrelated and weakly stationary in their first and second moments. To this end we make a simplifying assumption that the resulting sequence is independent and identically distributed from an unknown distribution. We fit Normal Inverse Gaussian and hyperbolic distribu-
tion to the residuals, estimated parameters of the two distribution are summarized in Table 3 and their corresponding kernel densities are compared in in Figure 1.

<table>
<thead>
<tr>
<th>Table 3: NIG and Hyperbolic parameter estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P500</td>
</tr>
<tr>
<td>NIG</td>
</tr>
<tr>
<td>α</td>
</tr>
<tr>
<td>β</td>
</tr>
<tr>
<td>δ</td>
</tr>
<tr>
<td>μ</td>
</tr>
</tbody>
</table>

4.2 Goodness of fit

Two sample test called Kolmogorov-Smirnov test (K-S test) was applied. Empirical CDF, an estimate of the underlying data is used to test the following hypothesis.

\[ H_0 : F_n(x) = F(x) \text{ for all } x \text{ versus } H_1 : F_n(x) \neq F(x) \text{ for some } x \]

where \( F_n(x) \) is the empirical cumulative probability estimated as \( F_n(x_i) = i/n \) for the ith smallest data value. \( F(x) \) is the theoretical cumulative distribution function evaluated at \( x \). The test statistic is given by

\[ D_n = \sup_x |F_n(x) - F(x)| \]

If the null hypothesis is true, then the theoretical distribution fits very well. If \( D_n \) is sufficiently large; the null hypothesis can be rejected. For the two filtered residual we computed \( D_n \) statistics and got \( D = 0.0163 \), \( pvalue = 1 \), for both indices (S&P500 and CAC40). This implies that the two sample sets seems to come from the same parent distribution. Attempts to test whether the unknown distributions came from either of the three proposed densities, was reported in Table 4. The graphical representation of Kernel densities in Figure 1, supports the claim that the two densities share the same parent distribution.

Judged on the distance between the empirical and theoretical distributions of the residuals, the picture changes slightly. The Kolmogorov-Smirnov test allows us to reject all Gaussian models. As the Kolmogorov-Smirnov statistic measures the uniform distance between two distribution functions, it might be of interest to
Figure 1: CAC40 filtered returns and S&P500 filtered returns density estimations

Table 4: Kolmogorov Smirnov distances

<table>
<thead>
<tr>
<th>$K_{Dist}$</th>
<th>NIG</th>
<th>HYP</th>
<th>NORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P500</td>
<td>0.4872</td>
<td>0.4872</td>
<td>0.7136</td>
</tr>
<tr>
<td>CAC40</td>
<td>0.5077</td>
<td>0.5077</td>
<td>0.7091</td>
</tr>
</tbody>
</table>

test the models ability to appropriately model leptokutic nature of log returns. If one suspects the data to be from a family of normal distribution, corresponding values for the Anderson - Darling distance which has its focus on the tails of the distribution will be valid. The following discrete version of the Anderson - Derling statistic $AD$, measuring the distance between the theoretical distribution function $F$ and the empirical distribution $\hat{F}$:

$$AD(F, \hat{F}) = \sup_{x \in \mathbb{R}} \frac{|F(x) - \hat{F}(x)|}{\sqrt{F(x)(1 - F(x))}}$$

From our study, it was clear that the data in question was non normal. Models based on the NIG and hyperbolic distribution which was used to provide a flexible description for empirically observed conditioned leptokurtic residuals seemed to fit the data.
5 Conclusions

Since log-Levy models fall short of explaining the auto correlation of the absolute return, we have observed that that the log return dynamics can be modelled by assuming three components (the mean function, the volatility function and the GARCH filtered residuals). Identification of the probability distributions can be estimated by class of distributions which can capture skewness and kurtosis for example normal inverse Gaussian or hyperbolic distributions. It is widely recognized that the key to developing successful strategies for managing risk and pricing assets is to parsimoniously describe the stochastic process governing the asset dynamics. This paper proposes general framework assumed to improve modeling returns of financial time series data. To this end, the proposed framework combine two major stylized facts of returns: changing variance and presence of excess kurtosis in filtered returns. The framework can be studied further by allowing different distributions under the objective measure $\mathbb{P}$. Empirical investigation indicate that the all the required model parameters can be estimated form from historical data and the filtered residual are non-normal. Further refinement of the model is left for future studies.

Acknowledgements

Comments from the Editor and the anonymous referee are acknowledged. Financial support from International Science Program(Sweden)/ EAUMP is greatly appreciated.

References


