Modeling Electricity Price Returns using Generalized Hyperbolic Distributions

F. Noyanim Nwobi[1]

Abstract

This paper compares the performances of five members of the Generalized Hyperbolic family of distributions (i.e., the Generalized Hyperbolic (GH), Hyperbolic (HYP), Normal Inverse Gaussian (NIG), Hyperbolic Skew Student's t (SSt) and Variance-Gamma (VG) distributions) alongside the Gaussian as benchmark in fitting log returns of an Electricity Futures Contract. Using log likelihood (LLH) function and Akaike Information (AIC) as criteria for selection, the GH and NIG outperformed other models, having 49.8% and 49.6% weight of evidence in their favour respectively for being the two models that give the best prediction of the log returns. However, simulation results show that GH is the most consistent among the candidate distributions especially in large sample situations. The tails behaviour of these distributions show that the SSt overestimates while the HYP and VG underestimate the probability of rare events in the electricity market at both tails. Results show that these distributions have substantial heavy tails.

Mathematics Subject Classification : 60G70**,** 62P05**,** 62F07, 62F10.

Keywords: Class of generalized hyperbolic distributions; Model selection; AIC; LLH; Tail behaviour; Heavy tails; Log returns.

1   Introduction

Participants in commodity markets strongly focus on unexpected price changes which are fundamentally determined by supply and demand imbalances. Of

---

[1] Department of Statistics, Imo State University, Owerri 460222, Nigeria
Email: fnnwobi@imsu.edu.ng

interest to these participants is the probability distribution of returns of the prices of a commodity for appropriate pricing regime and especially in measuring the expected risk in such investment.

The large price fluctuations frequently observed in energy markets lead to non-normal deviations from the long-term mean. Such fluctuations measured in terms of returns over short term intervals are known to be characterized by non-normality: more peaked and heavier tails. This violation of the normality hypothesis implied by geometric Brownian motion (see, for example, Fama [1], Mandelbrot [2]) necessitates the call for generalizations in modelling such large changes in prices.

Barndorff-Nielsen [3] found a good fit in Generalized Hyperbolic (GH) distribution to Danish stock returns. One appealing feature of the GH is that they combine the characteristics of the normal and stable distributions especially those of the Levy processes of hyperbolic type to offer more flexibility in modeling financial time series data. The Hyperbolic distribution, a subclass of the GH distribution, which has in addition exponentially decreasing tails, was independently suggested as distribution of German stock returns represented in the stock index DAX by Eberlein and Keller [4] and Küchler *et al.* [5]. In corroborating their study, Cont [6] added that for a parametric distribution to reproduce the properties of the empirical distribution, it must have at least four parameters, for example, location, scale, a parameter describing the decay of the tails and that for asymmetry parameter.

The Normal Inverse Gaussian (NIG) distribution was introduced in the last two decades or so see, e.g., Eberlein and Keller [4] and Prause [7] while modeling the time dynamics of financial markets by stochastic processes. Other 4-parameter distributions within this period are the Variance-Gamma (VG) by D. B. Madan and E. Seneta in 1990 (see, e.g., Madan and Seneta [8]) and the GH Student's

t-distribution (SSt) as limiting cases of the GH distribtion have recorded different but appreciable amounts of success in modelling returns of time series data. As we shall see, the tail behaviour of the GH distributions ranges from Gaussian through exponential tails to the Student's t-distribution power tails.

The problem of testing whether some given observations follow one of the listed probability distribution functions is quite old in the statistics literature. Among these models, the effect of choosing a wrong model was originally discussed by Cox [9] in general and it has been demonstrated it nicely by a real data example. Due to increasing applications of the heavy tail distributions, special attention is now given to the discrimination among the family of GH distributions. In this study we discriminate among the five members of the GH distribution family for a log returns series of an Electricity Futures Contract and their fits compared especially against the Gaussian model. Using the tails behaviour we identify a candidate distribution that best approximate the empirical distribution.

This paper is arranged as follows; a brief overview of the Generalized Hyperbolic distribution and its special and limiting cases are given in Section 2 while procedures for model selection are presented in Section 3. Implementation and simulation results are presented in Sections 4 and 5 respectively and we conclude in Section 6.

2   The Generalized Hyperbolic Family of Distributions

We start with an exposition of the univariate Generalized Hyperbolic (GH) Distribution introduced in the literature by Ole Barndorff-Nielsen in 1977 while modeling particle size from a diamond mine (see, e.g., Barndorff-Nielsen, [10]) and the subclasses which have been listed in Section 1. The distribution is well applied in economics particularly in the fields of modeling financial markets and risk management due to its semi-heavy tails. A random variable X is said to follow a GH distribution if it is of the form

$$X \sim GH\left(\lambda, \alpha, \beta, \delta, \mu\right)$$

where $x \in \mathbb{R}$ the parameters $\mu, \delta, \beta, \alpha$ respectively determine the location, scale, skewness and shape of the distribution while λ influences the kurtosis and characterizes the classification of the GH distribution.

2.1 The Generalized Hyperbolic Distribution

A random variable X is said to follow a Generalized Hyperbolic (GH) distribution if its probability density function is given by

$$f_{GH}(x; \alpha, \beta, \delta, \lambda, \mu) = \frac{\left(\gamma/\delta\right)^{\lambda}}{\sqrt{2\pi} K_{\lambda}\left(\delta\gamma\right)} \frac{K_{\lambda-\frac{1}{2}}\left[\alpha\sqrt{\delta^2 + \left(x-\mu\right)^2}\right]}{\left[\sqrt{\delta^2 + \left(x-\mu\right)^2 \big/ \alpha}\right]^{\frac{1}{2}-\lambda}} \exp\left\{\beta\left(x-\mu\right)\right\} \quad (1)$$

where $\gamma = \sqrt{\alpha^2 - \beta^2}$, the domain of variation of the parameters is $\mu, \alpha \in \mathbb{R}$, $\delta \geq 0$ and, $0 \leq |\beta| < \alpha$ while K_{λ} is the modified Bessel function of the third kind with index λ.

Special cases of the generalized hyperbolic distribution (see, e.g., Jørgensen [11], Barndorff-Nielsen and Stelzer [12]) are

  (i) when $\lambda = -1/2$, the GH specializes to the Normal Inverse Gaussian (NIG),

  (ii) when $\lambda = 1$, the GH becomes the Hyperbolic distribution and

  (iii) when $\alpha \to \infty, \beta = 0$ and $\delta = 2\sigma^2$, the GH converges to the normal
        distribution with mean μ and variance σ^2.

Definition 1 (*Modified Bessel Function of the Third Kind with Index λ*)
The integral representation of the modified Bessel function of the third kind with index λ can be found in Barndorff-Nielsen et al. [13] and Abramowitz and Stegun [14],

$$K_{\lambda}(x) = \frac{1}{2} \int_0^{\infty} y^{\lambda-1} \exp\left\{-\frac{x}{2}\left(y - y^{-1}\right)\right\} dy, \qquad x > 0. \quad (2)$$

The substitution $y = x\sqrt{\chi/\psi}$ can be used to obtain the following relation which allows one to bring the GH (1) into a closed-form expression

$$\int_0^\infty y^{\lambda-1} \exp\left\{-\frac{1}{2}\left[\frac{\chi}{y} + y\psi\right]\right\} dy = 2\left(\frac{\chi}{\psi}\right)^{\lambda/2} K_\lambda\left(\sqrt{\chi\psi}\right). \tag{3}$$

Asymptotic relations for small arguments x can be used for calculating the densities of special cases of the GH density as follows

$$K_\lambda\left(x\right) \sim \Gamma\left(\lambda\right) 2^{\lambda-1} x^{-\lambda} \quad \text{as } x \downarrow 0 \text{ and } \lambda > 0 \tag{4}$$

and

$$K_\lambda\left(x\right) \sim \Gamma\left(-\lambda\right) 2^{-\lambda-1} x^{\lambda} \quad \text{as } x \downarrow 0 \text{ and } \lambda < 0. \tag{5}$$

2.2 The Normal Inverse Gaussian

A random variable X follows a Normal Inverse Gaussian (NIG) distribution with parameter vector $(\alpha, \beta, \mu, \delta)$ if its probability density function is

$$f_{NIG}(x : \alpha, \beta, \mu, \delta) = \frac{\alpha\delta \exp\left[p\left(x\right)\right]}{\pi q\left(x\right)} K_1\left[\alpha q\left(x\right)\right] \tag{6}$$

where $p(x) = \delta\sqrt{\left(\alpha^2 - \beta^2\right)} + \beta\left(x - \mu\right)$, $q(r) = \sqrt{\left(x - \mu\right)^2 + \delta^2}$ and K_1 is the modified Bessel function of the third kind with order one (see e.g., Abramowitz and Stegun, [14]). Here $\mu, \beta \in \mathbb{R}$ but if $\beta < 0$, the NIG is negatively skewed; $\alpha \geq |\beta|$ measures the heaviness of the tails (shape of the distribution). The NIG is a very flexible member of the family of distributions enjoying the convolution property as shown in Kalemanova and Werner [15].

Property 1
The NIG is a mixture of normal and inverse Gaussian distributions.
Let

$$X|Y = \begin{cases} y \sim N(\mu + \beta y, y) \\ Y \sim IG(\delta\gamma, \gamma^2) \quad \text{with } \gamma := \sqrt{\alpha^2 - \beta^2} \end{cases} \tag{7}$$

then $X \sim NIG(\alpha, \beta, \mu, \delta)$ and is denoted by the density function

$$f_{NIG}(\alpha, \beta, \mu, \delta) = \int_0^\infty f_N(x; \mu + \beta y, y).f_{IG}(y; \delta\gamma, \gamma^2) dy. \tag{8}$$

Property 2

The NIG distribution is closed under convolution. In fact, it is the only member of the family of generalized hyperbolic distributions to have the property that for independent random variables, $X \sim NIG(\alpha, \beta, \mu_X, \delta_X)$ and $Y \sim NIG(\alpha, \beta, \mu_Y, \delta_Y)$, their sum is NIG distributed, that is,

$$X + Y \sim NIG(\alpha, \beta, \mu_X, \delta_X) * NIG(\alpha, \beta, \mu_Y, \delta_Y) = NIG(\alpha, \beta, \mu_X + \mu_Y, \delta_X + \delta_Y) \tag{9}$$

2.3 The Hyperbolic Distribution

The random variable X is said to have a Hyperbolic (HYP) distribution if its probability density function is given by

$$f_{HYP}(x; \alpha, \beta, \delta, \mu) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha\delta K_1(\delta\sqrt{\alpha^2 - \beta^2})} \exp\{-\alpha(u(x)) + \beta(x - \mu)\} \tag{10}$$

where $u(x) = \sqrt{(\delta^2 + (x - \mu)^2)}$ and $-\infty \le x \le \infty.$ The domain of variation of the parameters is $\mu \in \mathbb{R}, \delta > 0,$ and $0 \le |\beta| < \alpha.$

Aas and Haff [16] derived a probability density function as a limiting case of the GH distribution $(\lambda = -v/2$ and $\alpha \to |\beta|)$ in (1); they referred to it as GH skew Student's t-distribution. The main attraction of this distribution is that unlike any other member of the GH family, it has one tail determined by a polynomial and the other by exponential behaviour. In addition, it is almost as analytically

tractable as the NIG distribution. Therefore, the GH skew Student's t-distribution has one heavy and one semi-heavy tail.

2.4 The Skew Student's t-distribution

A random variable X is said to follow a GH skew Student's t-distribution (SSt) if its (Aas and Haff, [16]) probability density function is given by

$$f_{SSt}\left(x;\nu,\mu,\beta,\delta\right) = \begin{cases} \dfrac{2^{(1-\nu)/2}\delta^{\nu}\left|\beta\right|^{(\nu+1)/2}K_{(\nu+1)/2}\left(\left|\beta\right|u\left(x\right)\right)}{\Gamma\left(\dfrac{\nu}{2}\right)\sqrt{\pi}\left(u\left(x\right)\right)^{(\nu+1)/2}}\exp\left\{\beta\left(x-\mu\right)\right\} & \text{for } \beta \neq 0, \\[6mm] \dfrac{\Gamma\left(\dfrac{\nu+1}{2}\right)}{\delta\Gamma\left(\dfrac{\nu}{2}\right)\sqrt{\pi}}\left[1+\dfrac{\left(x-\mu\right)^{2}}{\delta^{2}}\right]^{-(\nu+1)/2} & \text{for } \beta = 0, \end{cases}$$

$$\tag{11}$$

where $u\left(x\right) = \sqrt{\left[\delta^{2}+\left(x-\mu\right)^{2}\right]}, \nu > 0, \sigma \equiv \delta > 0, \mu, \beta \in \mathbb{R}.$

It can be recognized that the density in (11) is that of a non-central (scaled) Student's t-distribution with ν degrees of freedom when $\beta = 0$.

2.5 The Variance-Gamma distribution

Let X be a continuous random variable. X is said to be distributed as the Variance-Gamma (VG) distribution if its probability density function of the form

$$f_{VG}\left(x;\alpha,\mu,\lambda,\beta\right) = \frac{\left(\alpha^{2}-\beta^{2}\right)^{\lambda}\left|x-\mu\right|^{\lambda-1/2}K_{\lambda-1/2}\left(\alpha\left|x-\mu\right|\right)}{\sqrt{\pi}\Gamma\left(\lambda\right)\left(2\alpha\right)^{\lambda-1/2}}\exp\left(\beta\left(x-\mu\right)\right), \tag{12}$$

where $-\infty < x > \infty$, μ (location parameter), α, β (asymmetry parameter) are real and $\lambda > 0$. $\Gamma\left(.\right)$ denotes the Gamma function and K_{λ} the Bessel function of the second kind.

The class of Variance-Gamma distributions is closed under convolution in the

following sense that if X_1 and X_2 are independent random variables that are variance-gamma distributed with the same values of the parameters α and β, but possibly different values of the other parameters, λ_1, μ_1 and λ_2, μ_2 respectively, then $X_1 + X_2$ is variance-gamma distributed with parameters $\alpha, \beta, \lambda_1 + \lambda_2$ and $\mu_1 + \mu_2$.

3   Procedures for Model Selection

Having reviewed the different competing models from the family of generalized hyperbolic distributions, we now outline different methods for choosing the best fitting model to our given dataset. Suppose there are two families, say, $\mathcal{F} = \left\{ f(x;\theta); \theta \in \mathbb{R}^p \right\}$ and $\mathcal{G} = \left\{ g(x;\varphi); \varphi \in \mathbb{R}^q \right\}$, the problem is to choose the correct family for a given dataset $\left\{ x_1, x_2, ..., x_n \right\}$. The methods we describe in the following Subsections are used for model discrimination in the Section 4.

3.1 Maximum Likelihood Criterion

Suppose a random variable X has a density function $f\left(x; \theta_1, \theta_2, ..., \theta_k\right)$ that depends on k parameters. Let $\hat{\theta}_i$ denote the maximum likelihood estimator (MLE) of θ_i for the likelihood function $\mathcal{L}\left(X, \theta\right) = \prod_{i=1}^{n} f\left(x_i; \theta_1, \theta_2, ..., \theta_k\right)$. Similarly let $\hat{\theta}_i'$ denote the MLE of θ_i' from another density function $g\left(x; \theta_1', \theta_2', ..., \theta_k'\right)$ with likelihood function $\mathcal{L}\left(X, \theta'\right)$. The maximum likelihood principle proposed in Cox [17] is a maximum likelihood ratio test procedure

$$\mathcal{T}\left(\hat{\theta}, \hat{\theta}'\right) = \sum_{i=1}^{n} \ln \left[\frac{f\left(x_i \mid \hat{\theta}\right)}{g\left(x_i \mid \hat{\theta}'\right)} \right], \tag{13}$$

where $\hat{\theta}$ and $\hat{\theta}'$ are maximum likelihood estimators of parameter vectors of

competing models. Because the estimators provide the *best* explanation of the observed data, we choose the density $\mathcal{F}$ if $\mathcal{T} > 0$, otherwise choose $\mathcal{G}$. The solution $\mathcal{T}$ is sometimes called the Cox's statistic. Lu *et al.* [18] observed that the statistic $\ln \mathcal{T}$ should be asymptotically normally distributed when properly normalized.

3.2 The Akaike Information Criterion

Suppose X is a continuous random variable as defined in Subsection 3.1 representing a model, say,

$$X = h(t,q) + \varepsilon \tag{14}$$

where *h* is a mathematical model, in our case, a probability density function; ε is a random error term that is independent and identically distributed with probability distribution such as the normal. This criterion known as the Akaike Information Criterion (AIC) is generally regarded as the first and still continues to be the most widely known model selection criterion because of its utilization of the relationship between the maximum likelihood and the Kullback-Leibler information. The motivation of this criteria is outlined in Akaike [19] and we summarize the procedure for implementation based on AIC value as follows:

(i)   The loss of information when a fitted model is used rather than the best approximating model is given by the AIC differences

$$\Delta_i = AIC_i - AIC_{\min} \tag{15}$$

where $AIC_{\min}$ is AIC value for the best model in the set.

(ii)   The likelihood of a model being useful in making inference concerning the relative strength of evidence for each of the models in the set is given by

$$\mathcal{L}\left(g_i \big| y\right) \propto \exp\left(-\tfrac{1}{2}\Delta_i\right). \tag{16}$$

(iii)   The *Akaike weight of evidence* in favour of model *i* being the best approximating model in the set is

$$w_i = \exp\left(-\tfrac{1}{2}\Delta_i\right) \bigg/ \sum_{r=1}^{R} \exp\left(-\tfrac{1}{2}\Delta_r\right) \qquad (17)$$

where R is the total number of models in the set. Readers interested in AIC are referred to Akaike [19] and Burnham and Anderson [20] for details.

3.3 The Normality Hypothesis

We discuss possible skewness in a model because it is fundamental to mainstream financial modeling, portfolio investment decisions, and in many statistical testing procedures relating to asset returns. Skewness is defined as follows

$$\gamma_1 = \mu^3 \big/ \mu_2^3. \qquad (21)$$

where $\mu^3 = E\left(x_i - \mu\right)^3$, E is the expectation operator, μ is the mean of random return variable x_i and $\mu_2 \equiv \sigma^2$ is the variance. For normal *distribution* $\gamma_1 = 0$, otherwise, the distribution is asymmetric. Skewness is positive when right hand tail is heavier and negative when left hand tail is heavier.

An undisputable exception from the classical asset returns normality assumption is that empirical returns distributions indicate substantial excess kurtosis. A large positive value for kurtosis indicates that the tails of the distribution are longer (heavier) than those of a normal distribution are while a negative value indicates shorter tails (becoming like those of a box-shaped uniform distribution). The kurtosis is defined as

$$\gamma_2 = \mu^4 \big/ \mu_2^2 - 3 = k - 3 \qquad (22)$$

where $\mu^4 = E\left(x_i - \mu\right)^4$. For normal distribution, the value of k is three. When the $\gamma_2 > 0$, the distribution is referred to as *leptokurtic* and called *platykurtic* if $\gamma_2 < 0$.

4   Implementation and simulation studies

The energy dataset we study is the Daily Electricity Prices for Pennsylvania State (PJMW) from January 01, 2002 to October 28, 2010 corresponding to 1,900 observations. If we let P_t represent electricity prices at time, t, we define log returns as $r_t = \log\left(P_t/P_{t-1}\right), t = 1, 2, \ldots, T$ and $r_t \in \mathbb{R}$. The log returns r_t is assumed to be i.i.d. random variable.

Implementation of these models is based on the R packages (e.g., *f*Basics, SkewHyperbolic, generalizedHyperbolic, HyperbolicDist and VarianceGamma) available from the Comprehensive R Archive Network (CRAN). With special statistical functions in these packages we implement density, cumulative distribution functions, quantiles and random seed generation. Other functions implement simulation for maximum likelihood estimates (mle) (especially the Nelder and Mead algorithm) of parameters and tests making descriptive statistics of data and comparative study of various classes of models possible. The mle of the parameter vector $\Phi = \left(\lambda, \alpha, \beta, \delta, \mu\right)$ from the given dataset for each distribution under study was used in the simulation study when $n = 50$, 100, 500, 1000, 5000 and 10000. This simulation was made possible through the algorithms provided in Atkinson [21] and Rydberg [22] and implemented in HyperbolicDist by exploiting the normal variance-mean mixture structure of the GH variables.

5   Estimation Results and Discussion

In this section we discuss the estimation results of each subclass of the Generalized Hyperbolic family of distributions reviewed in Section 2. The results including that of the Gaussian distribution are presented in Table 1. Using the likelihood criterion, the GH and the NIG have the highest LLH of -2497.952 and -2497.955, respectively. From Table 1 also, the GH and the NIG have the least

AIC value of 5003.904 and 5003.910 with Akaike weight of evidence of 0.4978 (or 49.8%) and 0.4963 (or 49.6%) respectively for being the best fitting models. Only GH and NIG out of the six competing models took up 99.4% weight of evidence for fitting the returns dataset. To discriminate between GH and NIG or to what extent GH is better than NIG we resort to evidence ratio (ER) $w_{NIG}/w_{GH} = 0.4978/0.4963 = 1.0030$, an insignificant difference, which shows that any of the two models is good enough in fitting the dataset among other candidate models. Use of the rule of the thumb given in Burnham and Anderson [20], a $\Delta_i < 2$ suggest substantial evidence for model *i,* values within the interval $3 \leq \Delta_i \leq 7$ indicate that the model has considerably less support whereas a

Table 1: Parameter* estimation and Model selection criteria

Parameter	Model					
	GH	HYP	NIG	SSt	VG	GAUSS
λ	-0.4631	1.0000	-0.5000	-	1.0270	-
α	0.7418	1.4871	0.7239	2.9923	0.9736	-
β	0.0570	0.0771	0.0566	0.0412	0.0699	-
δ/σ	0.7061	0.1072	0.7199	1.0947	0.9690	1.0002
μ	-0.0567	-0.0717	-0.0564	-0.0450	-0.0695	-0.0000
LLH	-2497.952	-2504.520	-2497.955	-2502.537	-2505.855	-2694.396
AIC	5003.904	5017.040	5003.910	5013.073	5019.710	5392.792
Δ_i	0.000	13.136	0.006	9.170	15.806	388.888
w_i	0.4978	0.0007	0.4963	0.0051	0.0002	0.0000

$*\, \alpha \rightarrow$ Shape; $\beta \rightarrow$ skewness/asymmetry; $\delta \rightarrow$ scale/spread; $\mu \rightarrow$ location

$\Delta_i > 10$ indicates that the model is very unlikely to fit the data well. With this rule, only GH and NIG qualify and other models especially GAUSS have no support in fitting the return series of electricity price contract.

The tail behaviours of the six models are highlighted in Figure 1. The duo of GH and the NIG are indistinguishable in their tail behaviours fitting the empirical data best in both tails. The VG and the HYP underestimates the probability distribution of both the left and right tail while the SSt overestimates the probability function.
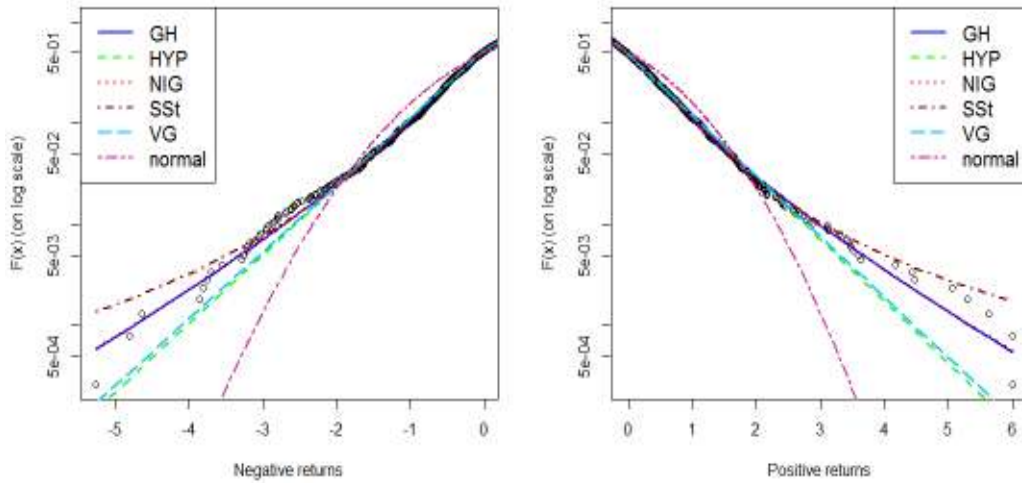
Figure1: Tail plots for returns: left tail (left panel) and right tail (right panel)

This implies that better events in the electricity market tend to occur more often than the HYP and VG distributions predict. Similarly, such events occur less often than the SSt distribution predicts. The Gaussian distribution overstates the probability distribution on both sides within the ranges (-2, 0) on the left tail and (0, 2) on the right tail (for the negative and positive returns respectively) and thereafter systematically understates the probabilities on both tails. The Gaussian distribution is therefore not a good predictor of the returns series of electricity futures prices.

The shape parameter which controls the decay in the tails is $\alpha \in (0,2]$, sometimes called the tail exponent. From Table 1 we infer that these five models belong to the class of heavy tail distributions. We defined a leptokurtic distribution (see equation (22)) as any distribution whose kurtosis is greater than zero, and by implication, all these distributions we compare in Table 3 have heavy tails.

The result of the simulation study is shown in Table 2. Again using the LLH and the AIC as criteria for selection, the GH is clearly the most consistent in the fit of the data under varying sample sizes especially when the sample size is large. This

study confirms HYP and VG to be out of contention as a model for this type of data set.

Table 2: Selection criteria based upon simulation results

n	Criteria	GH	HYP	NIG	SSt	VG	Gauss
50	LLH	-76.1753	-60.4157	-47.8386	-66.0349	-74.9982	-62.6272
	AIC	160.3506	128.8314	103.6772	140.0698	157.9964	133.2544
	w_i	0.0	0.0	**1.0**	0.0	0.0	0.0
100	LLH	-122.6249	-128.0804	-128.5928	-137.1527	-129.9325	-149.7984
	AIC	253.2498	264.1608	265.1856	282.3054	267.865	303.5968
	w_i	**0.9926**	0.0042	0.0025	0.0	0.0007	0.0
500	LLH	-640.4953	-704.5216	-663.4552	-634.9737	-666.7313	-726.2924
	AIC	1288.991	1417.043	1334.91	1277.947	1341.463	1456.585
	w_i	0.0040	0.0	0.0	**0.9960**	0.0	0.0
1000	LLH	-1222.445	-1289.924	-1346.708	-1316.855	-1338.404	-1390.847
	AIC	2452.89	2587.848	2701.416	2641.71	2684.808	2785.694
	w_i	**1.0**	0.0	0.0	0.0	0.0	0.0
5000	LLH	-6394.304	-6524.074	-6529.522	-6734.06	-6600.381	-7027.758
	AIC	12796.61	13056.15	13067.04	13476.12	13208.76	14059.52
	w_i	**1.0**	0.0	0.0	0.0	0.0	0.0
10000	LLH	-12920.07	-12995.03	-13015.28	-13152.71	-13092.94	-14159.72
	AIC	25848.14	25998.06	26038.56	26313.42	26193.88	28323.44
	w_i	**1.0**	0.0	0.0	0.0	0.0	0.0

The empirical and the estimated central moments of the returns are presented in Table 3. The estimated values of the main statistical indicators used in financial theory to access the returns distribution (i.e., mean, variance, skewness and kurtosis) are compared with the empirical values (i.e., computed values of the indicators from data) as a benchmark. Values show that the GH and NIG have estimates of the four indicators very close to empirical values.

Table 3: Empirical and estimated centered moments of the returns

Model	Moments			
	Mean	Variance	Skewness	Kurtosis
Empirical	**0.0000**	**1.0003**	**0.3346**	**4.9209**
GH	9.29×10^{-05}	1.0025	0.3213	4.7995
HYP	1.92×10^{-05}	0.9351	0.2123	2.8983
NIG	6.01×10^{-05}	1.0037	0.3254	5.9155
SSt	0.0048	0.4000	0.0404	1.2016
VG	4.00×10^{-04}	0.9437	0.2098	5.9502
Gauss	-0.0000	1.0002	0.0000	0.0000

6   Conclusion

In order to analyze energy futures prices of the deregulated Pennsylvania electricity market, we assumed that the log return series of the prices are driven by Levy process of the generalized hyperbolic type. We compared five members of this family (the generalized hyperbolic (GH), hyperbolic (HYP), normal inverse Gaussian (NIG), variance-gamma (VG) and hyperbolic skew Student t (SSt) distributions) along with the normal distribution as the benchmark.

Comparisons based upon the likelihood function (equation (13)), the Akaike information criteria (AIC) and some statistical indicators (i.e., the first four central moments) as criteria for selection are given in Table 1. Results in the table show that GH and NIG control 99.4% weight of evidence for being best two models among the six candidate probability distribution functions in the family. The two models are, however, indistinguishable in fitting the dataset.

Results from simulation study fovour GH as the best candidate probability model at varying sample sizes especially in large sample size situations. These distributions under study have shown substantial evidence of being heavy tailed using the values of the shape parameter, α, kurtosis, γ_2 and by visual inspection of Figure 1.

References

[1] E. Fama, The behaviour of stock market prices. *Journal of Business.* vol. **38**, 1965, pp. 34–105.

[2] B. Mandelbrot, The variation of certain speculative prices. *Journal of Business,* vol. **36**, 1963, pp. 394–419.

[3] O. E. Barndorff-Nielsen, Gaussian-inverse Gaussian process and the modeling of stock returns. Paper presented at the second Workshop on *Stochastics and Finance* 1994 in Berlin, 1994.

[4] E. Eberlin and U. Keller, Hyperbolic distributions in finance. *Bernoulli,* vol. **1**, no. 3, 1995, pp. 281–299.

[5] U. Küchler, K. Neumann, M. Sørensen, and A. Stroller, Stock returns and hyperbolic distributions. Discussion paper 23, Sonderforschungsbereich 373, Humboldt-Universität zu Berlin. Presented at the 2nd workshop on *Stochastics and Finance* in Berlin, 1994.

[6] R. Cont, Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, vol. **1**, 2001, pp. 223–236.

[7] K. Prause, The generalized hyperbolic models: Estimation, financial derivatives and risk measurement. PhD Thesis, Mathematics Faculty, University of Freiburg, 1999.

[8] D. B. Madan and E. Seneta, The variance gamma (V.G.) model for share market returns. *Journal of Business*, vol. **63**, 1990, pp. 511–524.

[9] D. R. Cox, *Tests of separate families of hypotheses*. Proceedings of the Fourth Berkeley Symposium in Mathematical Statistics and Probability, Berkeley, University of California Press, Berkley, CA, 1961, pp. 105–123.

[10] O. E. Barndorff-Nielsen, Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London*, vol. A**353**, 1977, pp. 401–419.

[11] B. Jørgensen, *Statistical properties of the generalized inverse Gaussian distribution*. Springer, New York, 1982.

[12] O. E. Barndorff-Nielsen and R. Stelzer, Absolute moments of generalized hyperbolic distributions and approximate scaling of normal inverse Gaussian Lévy processes. *Scandinavian Journal of Statistics,* vol. **32**, 2005, pp. 617–637.

[13] O. E. Barndorff-Nielsen, J. Kent, and M. Sorensen, Normal variance-mean mixtures and z–distributions. *International Statistical Review.* vol. **50**, 1982, pp. 145–159.

[14] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions.* Dover Publications, New York, 1972.

[15] A. Kalemanova and R. Werner, A short note on the efficient implementation of the normal inverse Gaussian distribution. URL http://www.mathfinance.ma.tum.de/peppers/KalemanovaWerner_NoteOnImplementation.pdf , 2006.

[16] K. Aas and I. H. Haff, The generalized hyperbolic skew Student's *t*-distribution. *Journal of Financial Econometrics,* vol. **4,** no. 2, 2006, pp. 275–309.

[17] D. R. Cox, Further results on tests of separate families of hypothesis. *Journal of the Royal   Statistical Society*, vol. B**24**, 1962, pp. 406–424.

[18] C. Lu, R. Danzer and D. Fishcer, Fracture statistics of brittle materials: Weibull or normal distribution. *Physical Review E,* vol. **65**, 2002, pp. 067–102.

[19] H. Akaike, A new look at the statistical model identification. *IEEE Transaction on Automatic Control,* vol. **19**, 1974, pp. 716–723.

[20] K. P. Burnham, and D. R. Anderson, *Model selection and inference: a practical information-theoretic approach*. Second Edition, Springer-Verlag, New York, 2002.

[21] A. C. Atkinson, The simulation of generalized inverse Gaussian and hyperbolic random variables. *SIAM Journal of Scientific and Statistical Computing,* vol. **3**, no. 4, 1982, pp. 502–515.

[22] T.H. Rydberg, The normal inverse Gaussian Levy process: Simulation and approximation. *Communications in Statistics-Stochastic Models*, vol. **34**, 1997, pp. 887–910.