Statistical Papers manuscript No.
(will be inserted by the editor)

GENERALIZED TUKEY-TYPE DISTRIBUTIONS WITH APPLICATION TO FINANCIAL AND TELETRAFFIC DATA

Matthias Fischer[1]

Lange Gasse 20, D-90403 Nürnberg

Received: 01.09.2006 / Revised version: date

Abstract Constructing skew and heavy-tailed distributions by transforming a standard normal variable goes back to Tukey (1977) and was extended and formalized by Hoaglin (1983) and Martinez & Iglewicz (1984). Applications of Tukey's GH distribution family – which are composed by a skewness transformation G and a kurtosis transformation H – can be found, for instance, in financial, environmental or medical statistics. Recently, alternative transformations emerged in the literature. Rayner & MacGillivray (2002b) discuss the GK distributions, where Tukey's H-transformation is replaced by another kurtosis transformation K. Similarly, Fischer & Klein (2004) advocate the J-transformation which also produces heavy tails but – in contrast to Tukey's H-transformation – still guarantees the existence of *all* moments. Within this work we present a very general kurtosis transformation which nests H-, K- and J-transformation and, hence, permits to discriminate between them. Applications to financial and teletraffic data are given.

1 Introduction

Using the Gaussian distribution as statistical model for data sets is widespread, especially in practice. However, departure from normality seems to be more the rule than the exception. In order to construct skew and heavy-tailed distributions, Tukey (1977) suggested to transform a standard Gaussian variable Z with a specific non-linear transformation, the so-called family of GH-transformations – which is composed by a skewness transformation G and a kurtosis transformation H. The corresponding GH distribution has been successfully applied in financial, medical or environmental

Send offprint requests to:

statistics. Badrinath & Chatterjee (1988, 1991) did pioneer work applying it to stock returns of the New York stock exchange. Mills (1995) applied them to FTSE index returns, Fischer, Horn & Klein (2006) to returns of aluminium and zinc, and Dutta & Babbel (2003) as distributional model for US-Dollar London Inter Bank Offer Rates. Furthermore, Dutta & Babbel (2005) and Tunaru, Kadam & Albota (2005) deal with derivative pricing formulas under GH distributions. In contrast, Dupuis (2004) and Dupuis & Field (2004) adress the problem of modelling extreme wind speeds and conclude that the GH distribution is "at least as effective as the more classical generalized extreme value distribution". Although GH distributions dominate both theory and application, different competitive transformations are available in the relevant literature: Rayner & MacGillivray (2002b) discuss the GK distributions, where Tukey's H-transformation is replaced by another kurtosis transformation K. Similarly, Fischer & Klein (2004) advocate the J-transformation which also produces heavy tails but – in contrast to Tukey's H-transformation – still guarantees the existence of *all* moments. Within this work we exclusively focus on kurtosis transformations and present a very general one (in section 2) which nests H-, K- and J-transformation and allows to discriminate between them. Properties of the corresponding distributions are dedicated to section 3, whereas empirical application can be found in section 4.

To be more precise, Tukey (1977) suggested to transform a Gaussian random variable Z by means of certain transformations $T(z)$ via

$$Y = Z \cdot T(Z)^{\theta}, \qquad \theta \in \mathbb{R}. \tag{1}$$

In order to make sure that the tails of the distribution of Y are heavier than those Z (which we focus on within this work), $T(z)$ has to be positive, symmetric and strictly monotone increasing for $z \geq 0$ (see, e.g. Hoaglin, 1983). Moreover, we restrict ourselves on non-negative parameters $\theta \geq 0$, where no transformation takes place if $\theta = 0$ (i.e. Z and Y coincide) and positive values of θ produce positive tail elongation. Originally, Tukey's suggests the H-transformation[1]

$$T_H(z) \equiv \exp(z^2) \tag{2}$$

which guarantees the existence of moments up to order $1/(2\theta)$ which in turn coincides with the asymptotic tail index of Y (see Proposition 1 in Morgenthaler & Tukey, 2000). Alternative transformations with existing moments but still heavy tails followed up by MacGillivray & Cannon's (1997) K-transformation (see also Rayner & MacGillivray, 2002a, 2002b)

$$T_K(z) \equiv 1 + z^2 \tag{3}$$

and by Fischer & Klein (2004) who discussed the J-transformation

$$T_J(z) \equiv \cosh(z) = 0.5\left(e^x + e^{-x}\right). \tag{4}$$

[1] Note that Tukey (1977) originally considered $T_H(z) = \exp(0.5z^2)$, instead.

Until Klein & Fischer (2006), no "superstructure" was available for these transformations which includes (2), (3) and (4) as special case. One possible superstructure is the "*power series representation*"

$$T_A(z) \equiv \sum_{i=0}^{\infty} a_i z^{2i} \tag{5}$$

for certain weights $a_i \in \mathbb{R}$, $i \geq 0$ which guarantee that the power series has a finite limit. In particular, Tukey's H-transformation is obtained for $a_i = 1/i!$, $i \in \mathbb{N}$, equation (3) is recovered if $a_0 = 1$, $a_1 = 1$ and $a_i = 0$, $i > 1$ and equation (4) setting $a_i = 1/(2i!)$.

2 A general kurtosis transformation

It was the power series representation of the exponential function, i.e.

$$\exp(z) = 1 + \sum_{i=1}^{\infty} \frac{z^i}{i!}$$

which motivated formula (5). Unfortunately, this formula is not very operational because an infinite number of unknown parameters $a_0, a_1, \ldots, a_i, \ldots$ have to be estimated to reveal the "data-generating transformation". Another representation of the exponential function is given by

$$\exp(z) = \lim_{n\to\infty} (1 + z/n)^n.$$

The latter motivates a second general transformation − which we term as HJK-transformation, henceforth −

$$T_{HJK}(z;\beta,n,g) = \left(1 + \frac{(z^2+g)^\beta - g^\beta}{n}\right)^n, \qquad \beta > 0,\ g > 0,\ n \geq 1. \tag{6}$$

Obviously, $T_H(z) = T_{HJK}(z;1,\infty,g)$ and $T_K(z) = T_{HJK}(z;1,1,g)$, i.e. both K-transformation and H-transformation are included as special case. Moreover, we will demonstrate later that $T_J(z) \approx T_{HJK}(x;0.5,\infty,0.5)$. Hence, accounting for the law of parsimony, it seems to reasonable setting $g \equiv 0.5$ in equation (6). Above these well-known examples, a great variety of transformations is included, some of them are shown in figure 1, below.
Obviously, positiveness and symmetry of T_{HJK} are guaranteed. Moreover, under the above assumptions and for $z > 0$,

$$T'_{HJK}(z;\beta,n) = 2\,\beta\,z\,\left(1 + \frac{\left(z^2+0.5\right)^\beta - 0.5^\beta}{n}\right)^{n-1} \left(z^2+0.5\right)^{\beta-1} > 0$$

i.e. the HJK-transformation is strictly increasing. Applying

$$(1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i$$

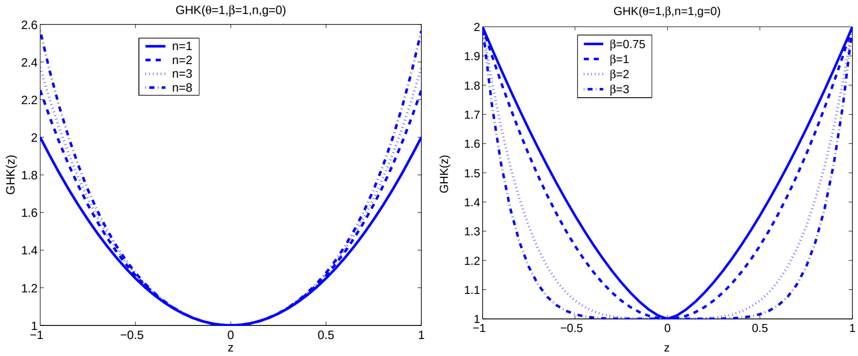

Fig. 1 Different GHK-transformations.

to equation (6),

$$T_{HJK}(z;\beta,n) = \sum_{i=0}^{n} \frac{\binom{n}{i}}{n^i} \Big((z^2+0.5)^\beta - 0.5^\beta \Big)^i. \tag{7}$$

For $\beta = 1$, the coefficients of the power series representation from (5) are

$$a_0 = 1, \quad a_i = \frac{\binom{n}{i}}{n^i}, \quad i = 1, \dots, n \text{ and } a_i = 0, \quad i > n.$$

If otherwise $\beta \approx 1$, an *approximate* power series representation can be obtained if a second order Taylor approximation at $z_0 = 0$ is applied to $(z^2 + 0.5)^\beta$, namely

$$(z^2+0.5)^\beta \approx 0.5^\beta + 2\beta\, 0.5^\beta z^2 \text{ and } T_{HJK}(z;\beta,n) \approx \sum_{i=0}^{n} \frac{\binom{n}{i}(2\beta 0.5^\beta)^i}{n^i} z^{2i}.$$

The approximate coefficients are then

$$a_0 = 1, \quad a_i = \frac{\binom{n}{i}(\beta\, 0.5^{\beta-1})^i}{n^i}, \quad i = 1, \dots, n \text{ and } a_i = 0, \quad i > n. \tag{8}$$

Within this work we focus on two special cases, the HK-transformation which nests both H- and K-transformation and the HJ-transformation which includes the H-transformation and closely approximates the J-transformation (as it will be shown later on).

2.1 The HK-Transformation

Setting $\beta = 1$ in (6) we obtain the HK-transformation

$$T_{HK}(z;n) = \left(1 + \frac{z^2}{n}\right)^n, \qquad n \geq 1. \tag{9}$$

Obviously, $T_H(z) = T_{HK}(z;\infty)$ and $T_K(z) = T_{HK}(z;1)$. Furthermore,

$$T_{HK}(z;2) = 1 + z^2 + 0.25z^4.$$

At any case, the (exact) coefficients of the power series representation are given by

$$a_0 = 1, \quad a_i(n) = \frac{\binom{n}{i}}{n^i}, \quad i = 1, \dots, n \text{ and } a_i(n) = 0, \quad i > n.$$

Note that for $j \geq 1$,

$$a_i(n) \leq a_i(n+j), \quad i = 1, \dots, \infty \quad \text{and} \quad \lim_{n\to\infty} a_i(n) = \frac{1}{i!}.$$

2.2 The HJ-Transformation

Similarly, for $n \to \infty$, (6) reduces to

$$T_{HJ}(z;\beta,g) = \exp\left((z^2+g)^\beta - g^\beta\right), \qquad \beta > 0,\ g \geq 0 \tag{10}$$

which we will call the HJ-transformation. Setting $\beta = 1$ in (10) recovers again the H-transformation of Tukey (1977). Next we derive an approximation for the J-transformation. For $\beta = 0.5$, equation (10) rewrites to

$$T_{HJ}(z;0.5,g) = \exp\left(\sqrt{z^2+g} - \sqrt{g}\right). \tag{11}$$

In addition, the J-transformation of Fischer & Klein (2004) can be approximated for sufficiently large z as follows

$$T_J(z) = 0.5\left(e^z + e^{-z}\right) \approx 0.5\exp(z) = \exp(z - \ln(2)). \tag{12}$$

Equating (11) and (12), we obtain $\sqrt{z^2+g} - \sqrt{g} = z - \ln(2)$ and we obtain an approximation of the J-transformation for both smaller and larger z using $T_{HJ}(z;0.5,0.5)$ with $g = (\ln(2))^2 \approx 0.5$. Table 1 illustrates how accurate the approximation works.

z	0	0.5	1	2	4	6	7
$T_J(z)$	1.00	1.17	1.68	4.11	28.48	204.43	549.26
$T_{HJ}(z,0.5,0.5)$	1.00	1.13	1.54	3.76	27.31	201.72	548.32
Percentual deviation	0.00	3.96	8.73	9.22	4.30	1.35	0.17

Table 1 Approximating the J-transformation.

Setting $\beta = 2$ and $g = 0.5$, (10) reduces to

$$T_{HJ}(z;\theta,\beta) = \exp\left(\theta\left[(z^4+z^2)\right]\right)$$

which resembles Morgenthaler and Tukey's (2000) HQ-transformation.

3 HJK(-NORMAL) DISTRIBUTIONS

Starting with a standard normal variable Z, we now focus on the distribution of the variable $Y = K(Z) = Z \cdot T(Z)^{\theta}$, where T is one of the Tukey-type transformations considered before. Applying standard methods of variable transformation, the density of Y requires the inverse transformation of T – which is rarely available in closed form – and is given by

$$f_Y(y) = \frac{f_Z(K^{-1}(y))}{K'(K^{-1}(y))} \quad \text{with} \quad K'(z) = T(z)^{\theta-1}(T(z) + \theta z T'(z)).$$

In contrast, the p-quantiles of Y admit a very simple representation, i.e

$$Q_Y(p) = K(Q_Z(p)),$$

where $Q_Z(p)$ denote the p-quantile of a standard Gaussian distribution. If a HJK-transformation is applied to Z, the resulting distribution will be termed as HJK-distribution, henceforth. As before, investigations concentrate on the HK-distribution and on the HJ-distribution.

3.1 The HK-Distribution

It is already known that all moments of the K-distribution – corresponding to the K-transformation – exist (see, e.g. Klein & Fischer, 2004). The same applies to HK-distribution as the next lemma will show. Hence, HK-distributions approximate the H-distributions (which are obtained as limiting case) but still guarantee for finite moments.

Lemma 1 *All moments of the HK-distribution exist for $\theta \geq 0$ and all $n \in \mathbb{N}$.*

Proof: Define $Y_{\theta,n} \equiv Z\left(1 + \frac{Z^2}{n}\right)^{\theta\, n}$, where Z is standard Gaussian. Obviously, for $\overline{\theta} = [\theta + 1]$, where $[a]$ is the smallest integer less than a, $E(Y_{\theta,n}) \leq E(Y_{\overline{\theta},n})$. Using that $E(Z^i) < \infty$ for all $i > 0$ and defining $\nu = k\overline{\theta}n \in \mathbb{N}$,

$$\begin{aligned} E\left(Y_{\overline{\theta},n}^k\right) &= E\left(Z^k\left(1 + \frac{Z^2}{n}\right)^{\nu}\right) \\ &= E\left(Z^k \sum_{i=0}^{\nu} \binom{\nu}{i} \frac{Z^{2i}}{n^i}\right) = \sum_{i=0}^{\nu} \binom{\nu}{i} \frac{E\left(Z^{2i+k}\right)}{n^i} < \infty \quad \square \end{aligned}$$

In order to demonstrate how close the HK-distribution approximates the H-distribution which appears as limit distribution, table 3 summarizes selected $(1-p)$-quantiles for different $n \in \mathbb{N}$. Choosing $N = 2000$ seems to be advisable in order to obtain a satisfying approximation in the tails.

p	1/4	1/8	1/32	1/64	1/256	1/512	1/1024
$n = 1$	1.0382	1.0880	1.1615	1.1888	1.2323	1.2502	1.2662
$n = 2$	1.0418	1.1069	1.2229	1.2712	1.3532	1.3886	1.4211
$n = 5$	1.0445	1.1246	1.3015	1.3885	1.5541	1.6326	1.7084
$n = 10$	1.0455	1.1323	1.3470	1.4639	1.7076	1.8327	1.9593
$n = 50$	1.0463	1.1395	1.3986	1.5584	1.9383	2.1602	2.4051
$n = 100$	1.0464	1.1405	1.4065	1.5738	1.9812	2.2252	2.4994
$n = 500$	1.0465	1.1413	1.4131	1.5869	2.0191	2.2838	2.5863
$n = 2000$	1.0465	1.1414	1.4144	1.5894	2.0266	2.2955	2.6039
$n = \infty$	1.0465	1.1415	1.4148	1.5903	2.0291	2.2995	2.6099

Table 2 $(1-p)$-quantiles of the HK-distribution ($\theta = 0.1$).

Van Zwet (1964) introduces a partial kurtosis ordering on the set of all symmetric, continuous and strictly monotone increasing distributions. In this concept, a symmetric distribution F has less kurtosis than a symmetric distribution G ($F \preceq_S G$), if $G^{-1}(F(x))$ is convex for $x > F^{-1}(0.5)$, where F^{-1} and G^{-1} denote the inverse cdf (or quantile function) of F and G, respectively. The next lemma shows that the parameter θ is a kurtosis parameter (in the sense of van Zwet) for every member of the HK-distribution family, i.e. for n fixed.

Lemma 2 *Assume that $n \in \mathbb{N}$ is fixed. The parameter $\theta \geq 1/(2n)$ of the HK-distribution is actually a kurtosis parameter, i.e. increasing θ corresponds to higher kurtosis and vice versa.*

Proof: Assume that $\theta_2 \geq \theta_1 \geq 1/(2n) > 0$. According to Theorem 1 of Klein & Fischer (2006), it suffices to verify that

$$\frac{T'_{HK}(z;\theta_2,n)}{T'_{HK}(z;\theta_1,n)} \geq \frac{T_{HK}(z;\theta_2,n)}{T_{HK}(z;\theta_1,n)} \quad \text{and} \quad \frac{T''_{HK}(z;\theta_2,n)}{T''_{HK}(z;\theta_1,n)} \geq \frac{T'_{HK}(z;\theta_2,n)}{T'_{HK}(z;\theta_1,n)}.$$

This, however, follows directly from (10),

$$T'_{HK}(z;\theta,n) = 2\left(\frac{n+z^2}{n}\right)^{\theta n} \frac{\theta z n}{(n+z^2)} \quad \text{and}$$

$$T''_{HK}(z;\theta,n) = 2\theta\, n\left(\frac{n+z^2}{n}\right)^{\theta n} \cdot \frac{z^2(2\,\theta\, n-1)+n}{(n+z^2)^2} \geq 0 \quad \text{if} \quad \theta \geq 1/(2n). \quad \square$$

Similarly, $n \geq 1/(2\theta)$ is a kurtosis parameter for fixed $\theta \geq 0$, too.

Lemma 3 *Assume that $\theta \geq 0$ is fixed. The parameter $n \geq 1/(2\theta)$ of the HK-distribution is actually a kurtosis parameter, i.e. increasing n corresponds to higher kurtosis and vice versa.*

Proof: Assume that $n_2 \geq n_1$. Again, it suffices to verify that

$$\frac{T'_{HK}(z;\theta,n_2)}{T'_{HK}(z;\theta,n_1)} \geq \frac{T_{HK}(z;\theta,n_2)}{T_{HK}(z;\theta,n_1)} \quad \text{and} \quad \frac{T''_{HK}(z;\theta,n_2)}{T''_{HK}(z;\theta,n_1)} \geq \frac{T'_{HK}(z;\theta,n_2)}{T'_{HK}(z;\theta,n_1)}.$$

Some straightforward reformulations show that this is equivalent to

$$\frac{n_2}{n_1} \geq \frac{n_2+z^2}{n_1+z^2} \quad \text{and} \quad \frac{z^2(2\theta n_2-1)+n_2}{z^2(2\theta n_1-1)+n_1} \geq \frac{n_2+z^2}{n_1+z^2}.$$

Whereas the first inequality is obvious, the second is valid for $\theta \geq 1/(2n_1)$, i.e. if both transformations are convex. □

Finally, different HK-distributions are plotted in figure 2, below.

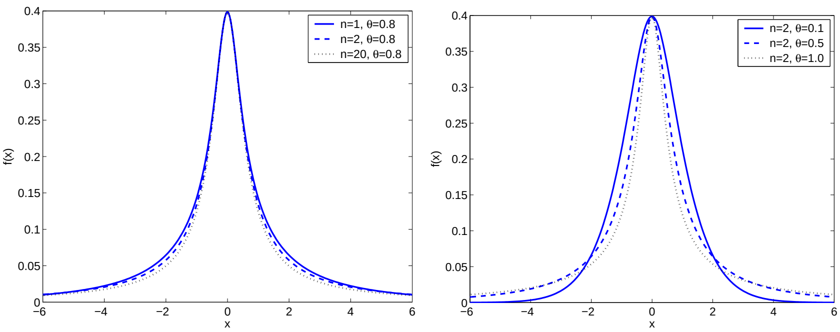

Fig. 2 Different Different HK-distributions.

To sum up, the heavy-tailed HK-distribution family provides an approximation of Tukey's H-distribution but still guarantees that all moments exist. In particular, it can be used to decide whether a H-transformation or a K-transformation is appropriate.

3.2 The HJ-Distributions

Consider the HJ-transformation with $g = 0.5$, i.e.

$$T_{HJ}(z;\theta,\beta) = \exp\left(\theta\left[(z^2+0.5)^\beta - 0.5^\beta\right]\right).$$

It was already mentioned that $\beta = 1$ corresponds to Tukey's H-transformation. The corresponding distribution has finite moments up to order less than $1/(2\theta)$ if Z is Gaussian. For $Y = Z \cdot T_{HJ}(Z;\theta,\beta)$ we obtain

$$\begin{aligned} E(Y^k) &= \int_{-\infty}^{\infty} z^k \exp\left(\theta k\left[(z^2+0.5)^\beta - 0.5^\beta\right]\right) \frac{1}{\sqrt{2\pi}} \exp(-0.5z^2)dz \\ &= \exp\left(-0.5^\beta \theta k\right) \int_{-\infty}^{\infty} z^k \frac{1}{\sqrt{2\pi}} \exp\left(\theta k(z^2+0.5)^\beta - 0.5z^2\right) dz \quad (13) \end{aligned}$$

Setting $\beta = 1$, we see that

$$E(Y^k) = \int_{-\infty}^{\infty} z^k \frac{1}{\sqrt{2\pi}} \exp\left(-0.5z^2(1-2\theta k)\right)\, dz < \infty$$

for $1 - 2\theta k > 1$, i.e. for $k < 1/(2\theta)$. Similar considerations – based on (13) – can be applied to show that no moments exist for $\beta > 1$ and all moments exist for $\beta < 1$. This explains (see table 3, below) why tails of the H-distribution ($\beta = 1$) are sufficiently approximated only for $\beta's$ very close to one.

p	1/4	1/8	1/32	1/64	1/256	1/512	1/1024
$\beta = 0.50$	1.0274	1.0664	1.1372	1.1688	1.2270	1.2541	1.2802
$\beta = 0.60$	1.0318	1.0805	1.1767	1.2227	1.3113	1.3545	1.3970
$\beta = 0.70$	1.0359	1.0949	1.2226	1.2878	1.4207	1.4884	1.5572
$\beta = 0.90$	1.0433	1.1254	1.3395	1.4663	1.7597	1.9278	2.1115
$\beta = 0.95$	1.0449	1.1333	1.3754	1.5247	1.8830	2.0956	2.3335
$\beta = 0.99$	1.0462	1.1398	1.4066	1.5765	1.9978	2.2554	2.5495
$\beta = 0.995$	1.0464	1.1407	1.4107	1.5834	2.0133	2.2772	2.5793
$\beta = 0.999$	1.0465	1.1413	1.4140	1.5889	2.0259	2.2950	2.6037
$\beta = \mathbf{1.00}$	1.0465	1.1415	1.4148	1.5903	2.0291	2.2995	2.6099

Table 3 $(1-p)$-Quantiles of the HJ-distribution ($\theta = 0.1$).

To sum up, the HJ-distribution family constitutes a very flexible distribution family which includes distributions for which all, some or no moments exist.

Lemma 4 *Assume that $\beta \geq 0.5$ is fixed. The parameter $\theta \geq 0$ of the HJ-distribution is actually a kurtosis parameter, i.e. increasing θ corresponds to higher kurtosis and vice versa.*

Proof: Its first derivative is

$$T'_{HJ}(z;\theta,\beta) = 2\beta\theta\ z\ \left(z^2+0.5\right)^{\beta-1} \exp\left(\theta\ \left(\left(z^2+0.5\right)^{\beta} + 0.5^{\beta}\right)\right) \quad (14)$$

which is monotone increasing for $z > 0$. Similarly,

$$T''_{HJ}(z;\theta,\beta) = 4\theta\beta\ \exp\left(\theta\ \left(z^2+0.5\right)^{\beta} + \theta\, 0.5^{\beta}\right) \left(z^2+0.5\right)^{\beta-2} \cdot R(z) \quad (15)$$

$$\text{with} \quad R(z) = \theta(z^2+0.5)^{\beta}\beta z^2 + z^2(\beta-0.5) + 0.25.$$

Thus, convexity of T_{HJ} is guaranteed at any case (independently of θ) if $\beta \geq 0.5$. Now assume that $\theta_2 \geq \theta_1 > 0$. In accordance to theorem 1 of Klein & Fischer (2006), it suffices to verify that

$$\frac{T'_{HJ}(z;\theta_2,\beta)}{T'_{HJ}(z;\theta_1,\beta)} \geq \frac{T_{HJ}(z;\theta_2,\beta)}{T_{HJ}(z;\theta_1,\beta)} \qquad \text{and} \qquad \frac{T''_{HJ}(z;\theta_2,\beta)}{T''_{HJ}(z;\theta_1,\beta)} \geq \frac{T'_{HJ}(z;\theta_2,\beta)}{T'_{HJ}(z;\theta_1,\beta)}.$$

The first inequality follows immediately after a few reformulations. The second inequality is equivalent to

$$\frac{\theta_2(z^2+0.5)^{\beta}\beta z^2 + z^2(\beta-0.5)+0.25}{\theta_1(z^2+0.5)^{\beta}\beta z^2 + z^2(\beta-0.5)+0.25} \geq 1.$$

At any case, this is true for $\beta \geq 0.5$ □

Lemma 5 *Assume that $\theta \geq 0$ is fixed. The parameter $\beta \geq 0.5$ of the HJ-distribution is actually a kurtosis parameter, i.e. increasing β corresponds to higher kurtosis and vice versa.*

Proof: Assume that $\beta_2 \geq \beta_1 \geq 0.5$. In accordance to the last lemma it suffices to verify that

$$\frac{T'_{HJ}(z;\theta,\beta_2)}{T'_{HJ}(z;\theta,\beta_1)} \geq \frac{T_{HJ}(z;\theta,\beta_2)}{T_{HJ}(z;\theta,\beta_1)} \quad \text{and} \quad \frac{T''_{HJ}(z;\theta,\beta_2)}{T''_{HJ}(z;\theta,\beta_1)} \geq \frac{T'_{HJ}(z;\theta,\beta_2)}{T'_{HJ}(z;\theta,\beta_1)}.$$

The first inequality equals

$$\beta_2(z^2+0.5)^{\beta_2} \geq \beta_1(z^2+0.5)^{\beta_1}$$

and follows from the monotonicity of the power function. The second inequality is equivalent to

$$\frac{\theta(z^2+0.5)^{\beta_2}\beta_2 z^2 + z^2(\beta_2-0.5)+0.25}{\theta(z^2+0.5)^{\beta_1}\beta_1 z^2 + z^2(\beta_1-0.5)+0.25} \geq 1.$$

and follows again if $\beta_2 \geq \beta_1 \geq 0.5$ □

Finally, different HJ-distributions are plotted in figure 3, below.

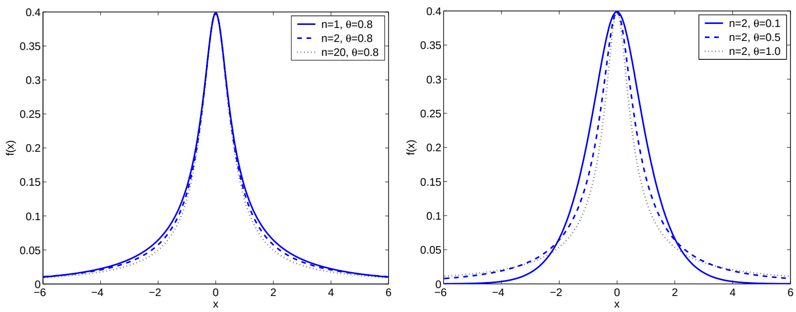

Fig. 3 Different Different HJ-distributions.

4 ESTIMATION OF HJK-DISTRIBUTIONS

Traditionally, quantile-based methods are applied to obtain estimates of the unknown parameters (see, e.g., Tukey, 1977 or Rayner & MacGillivray, 2002a). Due to the increasing computing power, maximum likelihood estimation (MLE) which had been thought intractable can now be tackled numerically. Rayner & MacGillivray (2002b) conducted a comprehensive study on MLE in the context of GH- and GK-distribution. Their results indicate that "sample sizes significantly larger than 100 should be used to obtain reliable estimates through maximum likelihood". Refering to Rayner & MacGillivray (2002b) for both theoretical and computational details, MLE maximizes the logarithm of the likelihood (as a function of the unknown parameters θ, β, n) for a simple random sample $y_1, \ldots, y_n$, given by

$$\mathcal{LL}(\theta, n, \beta; y_1, \ldots, y_n) = \sum_{i=1}^{n} \ln\left(\frac{f_Z(K^{-1}(y_i;\beta,n))}{K'(K^{-1}(y;\beta,n);\beta,n)}\right),$$

where K denote one of the Tukey-type transformation discussed above. Here numerical likelihood maximization was carried out using standard minimization routines in MATLAB which have the advantage of not requiring derivative information about the log-likelihood. The corresponding MATLAB code is available from the author by request.

5 APPLICATION TO HEAVY-TAILED DATA

At first, we focus on the continuously compounded returns (e.g. differences of consecutive log prices) of ALLIANZ AG over the period 1 January 1990 to 31 December 2003 (3485 observations). The (sample) mean of the log-returns (which are depicted in figure 4, below) is -0.00002 with a (sample) standard deviation of 0.0221. Moreover, the data set exhibits only a small amount of skewness (the skewness coefficient – measured by the third standardized moments – is given by by -0.069), whereas the kurtosis coefficient – in terms of the fourth standardized moments – is 5.362, reflecting the remarkable leptokurtosis.

The results for the ALLIANZ returns arising from maximum likelihood estimation of the parameters from different Tukey-type distributions are summarized in table 4, below.

Obviously, focussing on the log likelihood value $\mathcal{LL}$, the return data under consideration are closer to the H-distribution than to the K-distribution, but closer to the J-distribution than to the H-distribution. At any case, $\mathcal{LL}$ additionally increases both for HJ- and HK distributions. Altogether, the HJ-distribution family seems to be the best choice for the Allianz returns. This is also confirmed by the parameter estimators of the "super model" , i.e. of the HJK-distribution.

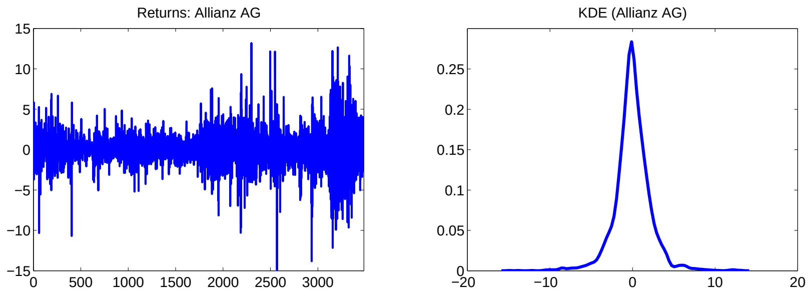

Fig. 4 Log returns and kernel density.

T	$\widehat{\mu}$	$\widehat{\delta}$	$\widehat{\theta}$	$\widehat{n}$	$\widehat{\beta}$	$\mathcal{LL}$
HJK	−0.0026	1.3808	0.2843	201.23	0.6208	−7310.7
H	−0.0084	1.465	0.2319	∞	**1.00**	−7315.7
HK	−0.0172	1.403	0.1750	5.50	**1.00**	−7311.2
K	−0.0264	1.285	0.3918	**1.00**	**1.00**	−7320.1
H	−0.0084	1.465	0.2319	∞	**1.00**	−7315.7
HJ	−0.0215	1.376	0.2975	∞	0.60	−7310.7
J	−0.0201	1.385	0.4059	∞	**0.50**	−7311.2

Table 4 ML estimation

The second data set is taken from the `BC-pAug89` data in the Internet Traffic Archive[2]. It measures the transferred bytes/sec within consecutive seconds. The trace `BC-pAug89` began at 11:25 on August 29, 1989, and ran for about 3143 seconds. The data are heavy-tailed featuring an excess kurtosis of 1.66. A further look on the time series and the corresponding kernel density estimation in figure 5 points out the significant skewness inherent to the internet traffic data (the third standardized moment is about 1.26).

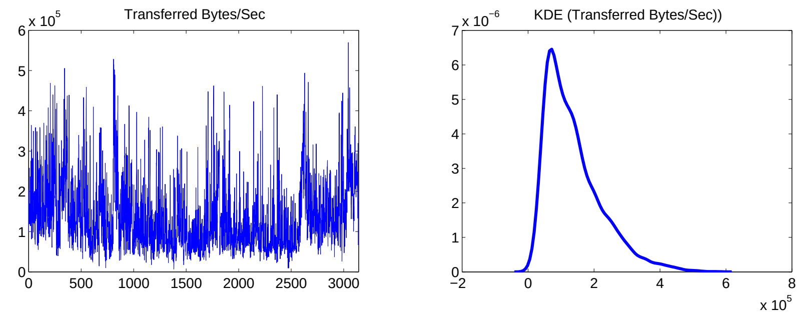

Fig. 5 Different HJ-distributions.

[2] `http://ita.ee.lbl.gov/html/contrib/BC.html`

For this reason, we also include a specific skewness transformation (for details on that transformations we refer to Rayner & MacGillivray, 2002a) to obtain the $G-HJK$-transformation

$$T_{G-HJK}(z) = \left(1 + 0.8 \cdot \frac{1-\exp(-gz)}{1+\exp(-gz)}\right) z \cdot T_{HJK}(z; \theta, n, \beta)$$

which reduces to the HJK-transformation in the symmetric case, i.e. for $g = 0$. Otherwise, the $G-HJK$ distributions are skewed to the right ($g > 0$) or skewed to the left ($g < 0$). Table 5 contains the ML-estimators and the corresponding log-likelihood values.

T	g	μ	δ	θ	n	β	$\mathcal{LL}$
$G-HJK$	0.7528	115690	77247	0.0078	199.9	1.52	-39546
$G-H$	0.7917	124130	88680	0.0348	∞	**1.00**	-39579
$G-HK$	0.7784	115681	77251	0.0241	199.8	**1.00**	-39549
$G-K$	0.7894	115440	72435	0.1123	**1.00**	**1.00**	-39568
$G-H$	0.7917	124130	88680	0.0348	∞	**1.00**	-39579
$G-HJ$	0.7528	115610	74931	0.0166	∞	1.15	-39549
$G-J$	0.7829	115600	74930	0.1008	∞	**0.50**	-39554

Table 5 ML estimation

Again, the J-transformation outperforms both H-transformation and K-transformation. Applying either the HK-transformation or the HJ-transformation instead of a "standard"-transformation increases the log-likelihood at any case. However, application of the HJK-transformation leads to a slight additional improvement of the log-likelihood-value. The estimation results suggest that $\beta \gg 1$, $n = \infty$ and $\theta \approx 0$ might be the best choice for the underlying data set.

6 SUMMARY

The construction of heavy-tailed distributions by transforming the standard normal variable goes back to Tukey (1977) who introduced the family of H-distributions or simply H-distributions derived from the H-transformation. Other transformations (e.g. K-transformation and J-transformation) followed up. Within this work we introduced and discussed the very general HJK-transformation and its corresponding distribution which nests H-, J- and K-transformations as (tractable) special case and allows to discriminate between them. Application to financial returns and internet traffic data is given.

References

1. Albota G, Kadam G, Tunaru R (2005) An Investigation of Parametric Risk Neutral Density Estimation. Working paper, City University of London.
2. Badrinath SG, Chatterjee S (1988) On measuring skewness and elongation in common stock return distributions: The case of the market index. Journal Business **61**:451-472.
3. Badrinath SG, Chatterjee S (1991) A data-analytic look at skewness and elongation in common stock return distributions. Journal of Business and Economic Statistics **9**:223-233.
4. Dutta KK, Babbel DF (2005) Extracting Probabilistic Information from the Price of Interest Rate Options: Tests of Distributional Assumptions. Journal of Business **78**(3), 841-870.
5. Dutta KK, Babbel DF (2003) On measuring skewness and kurtosis in short rate distributions: The case of the US dollar London Inter Bank Offer Rates. *Working paper 02-25* of the Wharton School, University of Pennsylvania.
6. Dupuis DJ, Field CA (2004) Large Wind Speeds: Modeling and Outlier Detection. Journal of Agricultural, Biological & Environmental Statistics **9**(1):105-121.
7. Field CA (2004) Using the *gh* distribution to model extreme wind speeds. Journal of Statistics, Planning and Inference **122**:15-22.
8. Fischer M, Horn A, Klein I (2006) Tukey-type Distributions in the Context of Financial Data. Communications in Statistics: Theory and Methods, accepted for publication.
9. Fischer M, Klein I (2004) Kurtosis Modelling by means of the J-Transformation. Allgemeines Statistisches Archiv **88**(1), 35-50.
10. Hoaglin DC (1983) Summarizing shape numerically: The $g-$ and $-h$ distributions, in: Hoaglin, DC, Mosteller F. & Tukey, JW (eds.): *Data analysis for tables, trends, and shapes*, Wiley, New York, 461-513.
11. Klein I, Fischer M (2006) Power Kurtosis Transformations: Definition, Properties and Ordering. Allgemeines Statistisches Archiv **90**(3), 395-402.
12. Martinez J, Iglewicz B (1984) Some Properties of the Tukey g and h Family of Distributions. Communications in Statistics: Theory and Methods **13**(3):353-369.
13. Mills TC (1995) Modelling skewness and kurtosis in the London Stock exchange FT-SE index return distribution. The Statistician **44**(3):323-332.
14. Morgenthaler S, Tukey JW (2000) Fitting quantiles: Doubling, HR, HQ, and HHH distributions. J. Comput. Graph. Stat. **9**(1):180-195.
15. Rayner GD, MacGillivray, HL (2002a) Weighted quantile-based estimation for a class of transformation distributions. Comput. Stat. Data An. **39**, 401-433.
16. Rayner GD, MacGillivray, HL (2002b) Numerical maximum likelihood estimation for the g-and-k and generalised g-and-h distributions.Statistics and Computing **12**, 57-75.
17. Tukey, JW (1977) *Exploratory data analysis*. Addison-Wesley, Reading, M.A.
18. Van Zwet, WR (1964) *Convex transformations of random variables*. Mathematical Centre Tracts No. 7, Mathematical Centre, Amsterdam.