International Journal of Mathematical Archive-4(5), 2013, 192-218
IJMA Available online through www.ijma.info ISSN 2229 – 5046

PORTFOLIO RANKING EFFICIENCY (I)

NORMAL VARIANCE GAMMA RETURNS

Werner Hürlimann*
Wolters Kluwer Financial Services, Seefeldstr. 69, CH-8008 Zürich

(Received on: 04-03-13; Revised & Accepted on: 06-05-13)

ABSTRACT

Equations of squared skewness and kurtosis as well as sharp inequalities between these quantities are derived for the normal bilateral gamma (NBG) convolution and the important normal variance gamma (NVG) sub-family. Application to portfolio selection with CARA utility is considered. With the NVG as test return distribution, it is analyzed whether a recent approximate ranking function with cubic mean-variance-skewness-kurtosis trade-off should be preferred to the original Gaussian ranking function with linear mean-variance trade-off or not. Based on an appropriate ranking efficiency measure and a simulation study, one notes, up to some exceptional cases, a systematic efficiency increase of the approximate ranking versus the Gaussian ranking. An empirical data analysis for eight different sets of returns from the Swiss Market and the Standard & Poors 500 stock indices, fitted to the NVG with the moment method, confirms the results from the simulation study. For this, full analytical solutions to the moment equations of the variance gamma and the normal variance gamma turn out to be very useful.

Mathematics Subject Classification: *60E15, 62E15, 62P05, 62P20, 91B16, 91G10.*

Keywords: *bilateral gamma, variance gamma, generalized gamma function, portfolio selection, ranking function, efficiency measure.*

1. INTRODUCTION

The four and five parameter normal Laplace (NL) and generalized normal Laplace (GNL) are two recent extensions of the normal and Laplace distributions (e.g. Reed and Jorgensen (2004), Reed (2006)). They are quite flexible for modelling purposes and have been used successfully in a number of different scientific contexts (e.g. Reed (2007/11), Reed and Pewsey (2009), Hürlimann (2008a/09/12)). The GNL is a re-parameterization of the normal variance gamma (NVG), which itself is a sub-family of the six parameter normal bilateral gamma (NBG) convolution (equations (2.4)-(2.6) in Section 2). The NVG contains the prominent and important variance gamma (VG), whose first version in financial research has been introduced by Madan and Seneta (1990) (see the Notes 2.2 for further information). The purpose of the present contribution is twofold. In the theoretical part, we aim a comprehensive understanding of the VG and NVG with regard to their skewness and (excess) kurtosis parameters. In particular, we are especially interested in their maximum domain of variation and the possibility to solve analytically the moment equations. The application part is directly related to the theoretical results. Due to a recent contribution by Di Pierro and Mosevich (2011), moment methods are particularly suited to analyze the portfolio selection problem within Financial Economics. For this we use equivalent ranking functions and define an appropriate ranking efficiency measure as explained in Appendix 1. Its practical use enables taking a decision about whether the recent approximate ranking function with cubic mean-variance-skewness-kurtosis trade-off by Di Pierro and Mosevich (2011) should be preferred to the original Gaussian ranking function with linear mean-variance trade-off by Lévy and Markowitz (1979) or not. It is important to remark that the developed theoretical results are not limited to this single application. The efficient and robust modelling of non-normality is also used in risk management and has gained importance due to the sub-prime and Euro crises as well as the new regulations in the finance industry like Basel III and Solvency II. A more detailed account of the theoretical and applied part follows.

Section 2 starts the theoretical part with a brief taxonomy of the NBG, which is defined as convolution of the normal and the bilateral gamma (BG). Two important members of the NBG are the VG and NVG, which play the main role in our application. A first main theoretical goal is a full analytical solution of the moment equations for the VG and NVG, which is presented in the Theorems 3.1 and 3.2 of Section 3. Preliminaries, which are required in the proofs, as well as important related results of independent interest, are summarized in the Appendix 2 and 3. For the BG and NBG we

Corresponding author: *Werner Hürlimann**
Wolters Kluwer Financial Services, Seefeldstr. 69, CH-8008 Zürich

derive equivalent scale parameterizations in terms of the kurtosis only, or in terms of the skewness and kurtosis (Theorems A2.1 and A2.4). We obtain sharp inequalities between skewness and kurtosis for them (Theorems A2.2 and A2.5). Moreover, parameterizations of the squared skewness and kurtosis parameters are displayed (Theorems A2.3 and Theorem A2.6), their extremal values are determined and their occurrences within the BG and NBG are discussed. Appendix 3 compares the skewness and kurtosis limits of the BG and NBG with other distributions. We show that the feasible domain of skewness and kurtosis for the BG contains the domain of the normal inverse Gaussian (NIG), which itself contains the domain of the generalized skew Student t (GST). Similarly, the domain of the BG encompasses the domain of the generalized t (GT) of Hansen (1994) provided the kurtosis exceeds the value 2.774. The Appendix 4 is devoted to special function representations of the BG density. They are used to derive analytical formulas for the densities and distribution functions of the VG and NVG. These are the required computational tools for a numerical evaluation of the goodness-of-fit statistics encountered in the empirical data analysis of Section 4.2.

The application to portfolio selection is presented in Section 4. It is based on the financial economics ranking efficiency measure defined and motivated in Appendix 1, Proposition A1.1. The investigation of this ranking efficiency measure for the NVG as test return distribution is illustrated at two different case studies. In the simulation study of Section 4.1 our calculations are based on monthly and quarterly equity return benchmark data. In Section 4.2 real-world equity return data sets from the Swiss Market and Standard & Poors 500 indices are fitted to the NVG return distribution and their ranking efficiency measures are calculated and compared. In the simulation study, we note, up to some exceptional cases, a systematic efficiency increase of the approximate ranking versus the Gaussian ranking. The empirical data analysis of Section 4.2 confirms this behaviour. To conclude, our statistical analysis shows that the approximate ranking function with cubic mean-variance-skewness-kurtosis trade-off (A1.1) should be preferred to the original linear mean-variance trade-off (A1.2), at least for the NVG test return distribution and up to pathological cases, which are not expected to occur in most of the current financial markets.

2. TAXONOMY OF THE NORMAL BILATERAL GAMMA CONVOLUTION AND SUB-FAMILIES

The *normal bilateral gamma* (NBG) random variable is defined to be the convolution of a normal and a bilateral gamma (BG) random variable. Following Küchler and Tappe (2008a/b)) for the BG, a six parameter NBG random variable takes the form

$$X = \upsilon + \tau \cdot Z + \alpha^{-1} \cdot G_1 - \beta^{-1} \cdot G_2 \sim NBG(\upsilon, \tau, \gamma, \alpha, \delta, \beta),\ \tau \geq 0,\ \gamma, \alpha, \delta, \beta > 0,\ -\infty < \nu < \infty, \quad (2.1)$$

with $Z \sim N(0,1)$ (standard normal), $G_1 \sim \Gamma(\gamma,1)$, $G_2 \sim \Gamma(\delta,1)$ (standardized gamma's with scale parameter 1), and (Z, G_1, G_2) independent. The NBG convolution includes a number of important and increasingly discussed families of distributions. We describe them by the choice of the parameters, specific naming and symbolic abbreviations (as usual μ, σ denote the mean and standard deviation, and for simplicity symmetric versions of the defined families are omitted):

Case 1: $\tau > 0$

$NVG(\upsilon, \tau, \rho, \alpha, \beta) = NBG(\upsilon, \tau, \gamma = \rho, \alpha, \delta = \rho, \beta)$: normal variance gamma (NVG)

$GNL(\upsilon, \tau, \rho, \alpha, \beta) = NVG(\upsilon\rho, \tau\sqrt{\rho}, \rho, \alpha, \beta)$: generalized normal Laplace (GNL)

$\ell NG(\upsilon, \tau, \rho, \beta) = \lim_{\alpha \to \infty} NVG(\upsilon, \tau, \rho, \alpha, \beta)$: left-tail normal gamma (ℓNG)

$rNG(\upsilon, \tau, \rho, \alpha) = \lim_{\beta \to \infty} NVG(\upsilon, \tau, \rho, \alpha, \beta)$: right-tail normal gamma (rNG)

$NL(\upsilon, \tau, \alpha, \beta) = NVG(\upsilon, \tau, \rho = 1, \alpha, \beta)$: normal Laplace (NL)

$\ell NE(\upsilon, \tau, \beta) = \lim_{\alpha \to \infty} NL(\upsilon, \tau, \alpha, \beta) = \ell NG(\upsilon, \tau, \rho = 1, \beta)$: left-tail normal exponential (ℓNE)

$rNE(\upsilon, \tau, \alpha) = \lim_{\beta \to \infty} NL(\upsilon, \tau, \alpha, \beta) = rNG(\upsilon, \tau, \rho = 1, \alpha)$: right-tail normal exponential (rNE)

Notes 2.1: The GNL distribution, which is a re-parameterization of the NVG, has been introduced in Reed (2006) (equations (2.4)-(2.6) below). Its generalization to the NBG has been mentioned in Lishamol and José (2009), Section 2, without detailed study, however. The convolution of a normal and a skew Laplace, which defines the normal Laplace, has been introduced in Reed and Jorgensen (2004) and Reed (2006). It can also be viewed as a convolution of a normal and a bilateral exponential. Since the Laplace and normal distributions constitute Laplace's first and second law of errors (e.g. Kotz et al. (2001), Chap. 1), it is worthy to consider a convolution of the two error distributions for modelling purposes. A statistical test to discriminate between both laws has been designed by Kundu (2005). A probabilistic genesis of the NL distribution has been provided by Reed and Jorgensen (2004). Such a convolution arises naturally if a Brownian motion with normally distributed initial state is observed at an exponentially distributed random time. This so-called exponential time changed Brownian motion with initial random normal state has the following meaning in finance applications. If the logarithmic price of a financial asset is assumed to follow a Brownian motion, then its logarithmic price at the *time of the first trade* on a fixed future date could be expected to follow a distribution close to a normal Laplace (e.g. Reed (2006), p.5). Similarly, a standardized gamma time changed Brownian motion with initial random normal state leads to a normal variance gamma distribution. The limiting cases (rNG) and (rNE) are simply convolutions of a normal with a gamma (NG) and an exponential (NE) respectively. We note that the latter two distributions are popular within the context of biological statistics and informatics (e.g. Irizarry et al. (2003), Xie et al. (2009), Plancade et al. (2011)). Clearly, the limiting case $\alpha \to \infty, \beta \to \infty$ of the normal Laplace is a normal distribution. From a broader classification viewpoint the NBG belongs to the class of *extended generalized gamma convolutions* (EGGC) introduced by Thorin (1978) and studied in particular by Bondesson (1992) (e.g. Küchler and Tappe (2008b), Section 3).

Case 2: $\tau = 0$

$BG(\upsilon,\gamma,\alpha,\delta,\beta) = NBG(\upsilon,\tau=0,\gamma,\alpha,\delta,\beta)$: bilateral gamma (BG)

$VG(\upsilon,\rho,\alpha,\beta) = BG(\upsilon,\gamma=\rho,\alpha,\delta=\rho,\beta)$: variance gamma (VG)

$GskL(\nu,\rho,\alpha,\beta) = GNL(\upsilon,\tau=0,\rho,\alpha,\beta)$: generalized skew Laplace (GskL)

$GL(\upsilon,\rho,\alpha) = VG(\upsilon,\rho,\alpha,\beta=\alpha)$: generalized Laplace (GL)

$skL(\upsilon,\alpha,\beta) = VG(\upsilon,\rho=1,\alpha,\beta)$: skew Laplace (skL)

$\ell G(\upsilon,\rho,\beta) = \ell NG(\upsilon,\tau=0,\rho,\beta)$: left-tail gamma (ℓG)

$rG(\upsilon,\rho,\alpha) = rNG(\upsilon,\tau=0,\rho,\alpha)$: right-tail gamma (rG)

Notes 2.2: The BG appears in a lot of recent papers (e.g. Küchler and Tappe (2009), Kaishev (2010), Bellini and Mercuri (2012), etc.). The VG special case with equal shape parameters of the gamma distributions and without shift has been introduced in Madan and Seneta (1990) and extensively studied (e.g. Madan and Milne (1991), Madan *et al.* (1998), Madan (2001), Carr *et al.* (2002), Geman (2002), Fiorani (2004), Fu *et al.* (2006), Stein *et al.* (2007), Domenig and Vanini (2010), etc.). Clearly, as in the Notes 2.1, the generalized skew Laplace is the re-parameterization of the variance gamma given by $GskL(\upsilon,\rho,\alpha,\beta) = VG(\upsilon\rho,\rho,\alpha,\beta)$. In Kotz *et al.* (2001), Chap.4, it is called a generalized Laplace, and alternatively a *Bessel function* distribution and a *variance-gamma* distribution. The first terminology is unfortunate because the here defined generalized Laplace (GL) has been studied previously by Mathai (1993a/b) and Koponen (1995). We prefer the new naming because the GskL is a generalization of the skew Laplace. Though omitted, some symmetric versions of these families are very important, as for example the symmetric Laplace $sL(\mu,\sigma) = skL(\upsilon=\mu,\alpha=\sigma^{-1}\sqrt{2},\beta=\alpha)$. Special instances of the BG have also been used earlier in actuarial science. For example, the generalized Laplace $GL(\upsilon=0,\rho,\alpha)$ with vanishing location $\upsilon=0$ has been considered in Chan (1998), p.89.

The cumulant generating function (cgf) of the NBG random variable is given by

$$C_X(t) = \ln\left\{E\left[e^{tX}\right]\right\} = \upsilon\cdot t + \tfrac{1}{2}\tau^2 t^2 + \gamma\cdot\ln\left\{\frac{\alpha}{\alpha-t}\right\} + \delta\cdot\ln\left\{\frac{\beta}{\beta+t}\right\}, \quad -\beta < t < \alpha. \tag{2.2}$$

From it one obtains the mean, variance and higher order cumulants as

$$\begin{aligned} &\mu = \upsilon + \gamma\cdot\alpha^{-1} - \delta\cdot\beta^{-1}, \quad \sigma^2 = \tau^2 + \gamma\cdot\alpha^{-2} + \delta\cdot\beta^{-2}, \\ &\kappa_r = C_X^{(r)}(t=0) = (r-1)!\left(\gamma\cdot\alpha^{-r} + (-1)^r\delta\cdot\beta^{-r}\right), \quad r > 2. \end{aligned} \tag{2.3}$$

The BG and NBG satisfy a number of important properties. For example, the BG has a smooth (differentiable) unimodal probability density function (pdf) (e.g. Küchler and Tappe (2008a)). Since the NBG belongs to the EGGC class (Notes 2.1), it is automatically self-decomposable, and henceforth unimodal (Bondesson (1992), Sato (1999), Cor. 15.11). In case $\alpha=\beta, \gamma=\delta$ the pdf is symmetric and bell-shaped taking a position between the normal and the symmetric generalized Laplace $sGL(\mu,\rho,\alpha) = GL(\upsilon=\mu,\rho,\alpha)$. A characterization of the sGL, which directly generalizes the symmetric Laplace, is presented in Corollary A2.1. The parameters α, β determine the behaviour in the left and right tail respectively. The NBG family is infinitely divisible and closed under linear transformation. It is in general not closed under convolution but a lot of sub-families share this property (e.g. the GNL).

The cgf of a GNL random variable $Y \sim GNL(\upsilon,\tau,\rho,\alpha,\beta)$, is related to the normal Laplace special case $\gamma = \delta = 1$ of the cgf (2.2) as follows:

$$C_Y(t) = \rho \cdot C_X(t). \tag{2.4}$$

From this relationship one sees without difficulty that the GNL random variable takes the form

$$Y = \rho\upsilon + \tau \cdot \sqrt{\rho} \cdot Z + \alpha^{-1} \cdot G_1 - \beta^{-1} \cdot G_2, \tag{2.5}$$

where Z, G_1, G_2 are independent with $Z \sim N(0,1)$, and G_1, G_2 are standardized gamma with scale parameter 1 and shape parameter ρ, that is with pdf $g(z) = \Gamma(\rho)^{-1} z^{\rho-1} e^{-z}$. The mean, variance and higher order cumulants are similarly to (2.3) given by

$$\begin{aligned} &\mu = \rho \cdot \left(\upsilon + \alpha^{-1} - \beta^{-1}\right), \quad \sigma^2 = \rho \cdot \left(\tau^2 + \alpha^{-2} + \beta^{-2}\right), \\ &\kappa_r = C_Y^{(r)}(t=0) = \rho \cdot (r-1)!\left(\alpha^{-r} + (-1)^r \beta^{-r}\right), \quad r > 2. \end{aligned} \tag{2.6}$$

3. SOLVING THE MOMENT EQUATIONS FOR THE VG AND NVG

For mathematical analysis of the moment equations for the bilateral gamma $BG(\upsilon,\gamma,\alpha,\delta,\beta)$ it is convenient to rescale the scale parameters with the standard deviation using the transformations

$$\alpha = (a \cdot \sigma)^{-1}, \quad \beta = (b \cdot \sigma)^{-1}, \quad a, b \geq 0. \tag{3.1}$$

The parameter set (a,b) is called *scale parameter set* and the parameters themselves *scale parameters*. Instead of the skewness and (excess) kurtosis parameters S, K, we use the proportionally scaled quantities

$$c = \tfrac{1}{2} S, \quad d = \tfrac{1}{6} K, \tag{3.2}$$

such that the moment equations can be rewritten as

$$\begin{aligned} &\upsilon + (\gamma \cdot a - \delta \cdot b)\sigma = \mu, \quad \gamma \cdot a^2 + \delta \cdot b^2 = 1, \\ &\gamma \cdot a^3 - \delta \cdot b^3 = c, \quad \gamma \cdot a^4 + \delta \cdot b^4 = d. \end{aligned} \tag{3.3}$$

By abuse of notation the BG family and sub-families are rewritten in the changed parameter space as $BG(\upsilon,\gamma,a,\delta,b)$, $VG(\upsilon,\rho,a,b)$, etc. A main goal is a full analytical solution of the moment equations for the VG and NVG distributions. Preliminaries, which are required in the proofs, as well as important related results of independent interest, are summarized in the Appendix 2 and 3.

Theorem 3.1: (*VG moment equations*). Given is a feasible skewness and kurtosis pair (S,K) satisfying the inequality $S^2 \leq \frac{2}{3} K$. Then, there exists a unique and explicitly given variance gamma distribution $VG(\upsilon,\rho,a,b)$, which solves the equation of skewness and kurtosis. Its parameters are fully analytical and specified as follows.

Case 1: $1 < \omega = 3(1 - K^{-1} S^2) \leq 3$

$$\begin{aligned} &\upsilon = \mu + \rho(b-a)\sigma, \quad a = \sqrt{\frac{1+\xi}{2\rho}}, \quad b = \sqrt{\frac{1-\xi}{2\rho}}, \quad \xi = \operatorname{sgn}(S) \cdot \sqrt{1 - X^2}, \\ &\rho = 3(2 - X^2) \cdot K^{-1}, \quad X = \frac{1}{3}\left(\sqrt[3]{\psi(\omega) + \sqrt{\Delta(\omega)}} + \sqrt[3]{\psi(\omega) - \sqrt{\Delta(\omega)}} - \omega\right), \\ &\psi(\omega) = 27(\omega - 1) - \omega^3, \quad \Delta(\omega) = 27(\omega - 1)(\psi(\omega) - \omega^3). \end{aligned} \tag{3.4}$$

Case 2: $\omega = 1 \quad \Leftrightarrow \quad K^{-1}S^2 = \frac{2}{3}$

$$\upsilon = \mu + \rho(b-a)\sigma, \quad a = \sqrt{\frac{1+\xi}{2\rho}}, \quad b = \sqrt{\frac{1-\xi}{2\rho}}, \quad \xi = \mathrm{sgn}(S), \ \rho = 6.K^{-1} \tag{3.5}$$

Proof: First of all, setting $w = \rho K$ one observes that solving the equation in (A2.9) is equivalent to finding a solution $(w, \rho), \ w \in [3,6], \ \rho > 0$, to the pair of equations

$$w - 2 - \sqrt{\left(\frac{6-w}{3}\right)^3} = \rho S^2, \quad w = \rho K .$$

With the change of variables $w = 3(2 - X^2)$ this task is equivalent to finding a solution $(X \in [0,1], \rho > 0)$ to the system of 2 polynomial equations in 2 unknowns

$$4 - 3X^2 - X^3 = \rho S^2, \quad 3(2 - X^2) = \rho K .$$

Inserting the second into the first equation, it suffices to solve for given $\omega = 3(1 - K^{-1}S^2) \in [1,3]$ the cubic equation $X^3 + \omega X^2 - 2(\omega - 1) = 0$. If $\omega = 1$ (Case 2) the equation has a double zero at $X = 0$. Otherwise, the unique solution follows from Cardano's formula. The parameters are found through inspection of Theorem A2.1. ◊

Remark 3.1: The existence of a unique solution has also been shown in Ghysels and Wang (2011), Proposition 2.4. The new derivation is simpler and yields a closed-form solution.

To investigate the moment equations for the normal bilateral gamma $NBG(\upsilon, \tau, \gamma, \alpha, \delta, \beta)$ let us rescale the parameters using (3.1) and (3.2) and set further

$$\tau = \sigma \cdot \sqrt{1 - s^2}, \quad s \in (0,1]. \tag{3.6}$$

With this transformation the moment equations of the NBG can be written as

$$\upsilon + (\gamma \cdot a - \delta \cdot b)\sigma = \mu, \quad \gamma \cdot a^2 + \delta \cdot b^2 = s^2, \gamma \cdot a^3 - \delta \cdot b^3 = c, \quad \gamma \cdot a^4 + \delta \cdot b^4 = d. \tag{3.7}$$

Interpreted in terms of (3.7), the BG moment equations (3.3) can be viewed as the degenerate case $s = 1$ of the pencil of ellipses $\gamma \cdot a^2 + \delta \cdot b^2 = s^2$, $s \in (0,1]$. From a mathematical point of view the analysis can be reduced to that of the BG using the iterated scale transformation

$$a = \overline{a}s, \quad b = \overline{b}s, \quad c = \overline{c}s^3, \quad d = \overline{d}s^4 \quad s \in (0,1]. \tag{3.8}$$

Applied to (3.7) one obtains the two-fold scaled equations of variance, skewness and kurtosis

$$\gamma \cdot \overline{a}^2 + \delta \cdot \overline{b}^2 = 1, \quad \gamma \cdot \overline{a}^3 - \delta \cdot \overline{b}^3 = \overline{c}, \quad \gamma \cdot \overline{a}^4 + \delta \cdot \overline{b}^4 = \overline{d}. \tag{3.9}$$

The NBG system (3.9) has the same form as the BG system (3.3). Henceforth, a great part of the NBG analysis directly follows from the BG case (see Appendix 2 for details). Theorem 3.1 generalizes as follows to the NVG family.

Theorem 3.2: (*NVG moment equations*). Given is a skewness and kurtosis pair (S, K) satisfying the inequality $S^2 \le \frac{2}{3}K$. For each $s \in [|S| \cdot \sqrt{\frac{3}{2}/K}, 1]$ there exists a unique and explicitly given normal variance gamma distribution $NVG\left(\upsilon, \tau = \sigma\sqrt{1-s^2}, \rho, a, b\right)$, which solves the equation of skewness and kurtosis. Its parameters are fully analytical and specified as follows:

Case 1: $1 < \omega = 3(1 - s^{-2}K^{-1}S^2) \le 3$

$$\begin{aligned}
&\upsilon = \mu + \rho(b-a)\sigma, \quad a = s\sqrt{\frac{1+\bar{\xi}}{2\rho}}, \quad b = s\sqrt{\frac{1-\bar{\xi}}{2\rho}}, \quad \bar{\xi} = \mathrm{sgn}(S)\cdot\sqrt{1-X^2}, \\
&\rho = 3(2-X^2)\cdot s^4K^{-1}, \quad X = \frac{1}{3}\left(\sqrt[3]{\psi(\omega)+\sqrt{\Delta(\omega)}} + \sqrt[3]{\psi(\omega)-\sqrt{\Delta(\omega)}} - \omega\right), \\
&\psi(\omega) = 27(\omega-1) - \omega^3, \quad \Delta(\omega) = 27(\omega-1)(\psi(\omega)-\omega^3).
\end{aligned} \tag{3.10}$$

Case 2: $\omega = 1 \quad \Leftrightarrow \quad K^{-1}S^2 = \frac{2}{3}s^2$

$$\upsilon = \mu + \rho(b-a)\sigma, \quad a = s\sqrt{\frac{1+\bar{\xi}}{2\rho}}, \quad b = s\sqrt{\frac{1-\bar{\xi}}{2\rho}}, \quad \bar{\xi} = \mathrm{sgn}(S), \quad \rho = 6s^4K^{-1}. \tag{3.11}$$

Proof: First of all, setting $s^4 w = \rho K$ one observes that solving the equation in (A2.15) is equivalent to finding a solution $(w,\rho),\ w \in [3,6],\ \rho > 0$, to the pair of equations

$$w - 2 - \sqrt{\left(\frac{6-w}{3}\right)^3} = s^{-6}\rho S^2, \quad w = s^{-4}\rho K.$$

With the change of variables $w = 3(2 - X^2)$ this task is equivalent to finding a solution $(X \in [0,1], \rho > 0)$ to the system of 2 polynomial equations in 2 unknowns

$$4 - 3X^2 - X^3 = s^{-6}\rho S^2, \quad 3(2 - X^2) = s^{-4}\rho K. \tag{3.12}$$

Inserting the second into the first equation it suffices to solve for given $\omega = 3(1 - s^{-2}K^{-1}S^2) \in [1,3]$ the cubic equation $X^3 + \omega X^2 - 2(\omega - 1) = 0$. We conclude as in the proof of Theorem 3.1, where the parameters are found through inspection of Theorem A2.4. ◊

4. APPLICATION TO THE RANKING EFFICIENCY IN PORTFOLIO SELECTION

The examination of the ranking efficiency measure (A1.8) with the NVG as test return distribution is illustrated at two different case studies. In the simulation study of Section 4.1 our calculations are based on the monthly and quarterly equity return benchmark data provided on the website of Prof. Kenneth French over the period from July 1926 to December 2010 and also used in Skoulakis (2012), Section 3. In Section 4.2 real-world equity return data sets from the Swiss Market and Standard & Poors 500 indices are fitted to the NVG return distribution and their ranking efficiency measures are calculated and compared.

4.1. Simulation of ranking efficiency

Computational evaluation of (A1.8) requires formulas for the NVG test ranking function $R_*(p^{NVG})$ and the approximate ranking function $R_*^A(p)$. The skewness and kurtosis parameters are those of the NVG distribution. According to (A1.1) and Theorem A2.6 one has

$$\begin{aligned}
&R_*^A(p) = \mu - \frac{m\sigma^2}{2} + \frac{m^2\sigma^3 S}{6} - \frac{m^3\sigma^4 K}{720}, \quad K = w\cdot\rho^{-1}\cdot s^4, \\
&S^2 = \left[w - 2 - \sqrt{\left(\frac{6-w}{3}\right)^3}\right]\cdot\rho^{-1}\cdot s^6, \quad w \in [3,6],\ s \in (0,1]
\end{aligned} \tag{4.1}$$

The relationship $R_*(p^{NVG}) = 1 + CER(R_U(p^{NVG}))$ (equation (A1.1) and CER definition) for a CARA utility function $U(x) = -\exp(-mx)$ is determined by

$$R_U(p^{NVG}) = -\exp\left\{-m\nu + \tfrac{1}{2}m^2\tau^2\right\}\left[\frac{\alpha\beta}{(\alpha+m)(\beta-m)}\right]^{\rho}. \tag{4.2}$$

With the transformation $\alpha = (a\cdot\sigma)^{-1}$, $\beta = (b\cdot\sigma)^{-1}$ it follows that

$$R_*(p^{NVG}) = -\ln(-R_U(p^{NVG}))/m = \nu - \tfrac{1}{2}m\tau^2 + \frac{\rho}{m}\cdot\ln\{(1+ma\sigma)(1-mb\sigma)\}. \tag{4.3}$$

Moreover, using (3.10) one gets further

$$\begin{gathered} R_*(p^{NVG}) = \mu + \rho(b-a)\sigma - \tfrac{1}{2}m(1-s^2)\sigma^2 + \frac{\rho}{m}\cdot\ln\left\{(1+ma\sigma)(1-mb\sigma)\right\}, \\ a = s\sqrt{\frac{1+\bar{\xi}}{2\rho}},\quad b = s\sqrt{\frac{1+\bar{\xi}}{2\rho}},\quad \bar{\xi} = \operatorname{sgn}(S)\sqrt{\frac{w-3}{3}},\quad w\in[3,6],\quad s\in(0,1]. \end{gathered} \tag{4.4}$$

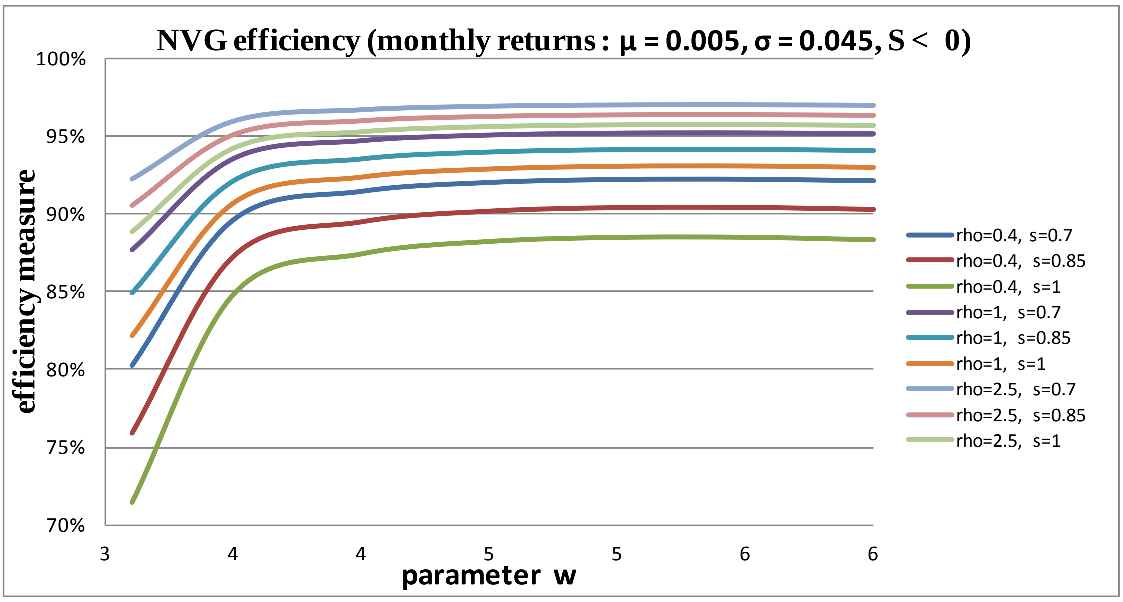
In the first case study, the equity return benchmark data comprises two sets of monthly and quarterly returns, whose percentage mean and standard deviation are chosen as follows:

monthly data: $(\mu,\sigma)\in\{\,(1,4.5), (1.5,6), (2,7.5), (2.5,9)\}$

quarterly data: $(\mu,\sigma)\in\{\,(3,7.5), (4.5,10), (6,12.5), (7.5,15)\}$

The qualitative impact of the skewness and kurtosis parameters is analyzed by varying the range of the parameters (ρ, s, w) and the sign of skewness in (4.1) and (4.4). The absolute risk aversion is first fixed at $m=2$, and the effect of its variation is mentioned later on. Numerical calculations reveal a systematic efficiency increase of the approximate ranking versus the Gaussian ranking by arbitrary sign of skewness over divers range of parameter values, e.g. $\rho \ge 0.12$, $s\in(0.001,1]$, $w\ge 3.2$, for the monthly returns, and $\rho\ge 0.35$, $s\in(0.001,1]$, $w\ge 3.2$, for the quarterly returns. In case the NVG is close to a symmetric distribution, i.e. $S\approx 0$ for $w\approx 3$, an efficiency decrease occurs. In general, the efficiency increase decreases with increasing values of (μ,σ), which implies lower bounds for ρ. It is possible to obtain parameter constellations for which the efficiency decreases with respect to the Gaussian ranking. For example, with the monthly returns $(\mu,\sigma)\in\{\,((2\%,7.5\%), (2.5\%,9\%)\}$, $\rho = 0.0016, s = 0.2, w = K\in[3,6], S\in[-2,0]$, the efficiency measure is always negative. To reject pathological cases with negative efficiency, a deeper statistical analysis of skewness and kurtosis is needed. The empirical data analysis of Section 4.2 suggests the approximate parameters $\mu = 0.5\%, \sigma = 4.5\%$ for the monthly returns of the Swiss Market and Standard & Poors 500 indices over periods of 24 and 63 years. The parameters (ρ, s) vary in the intervals $\rho\in[0.4,2.5], s\in[0.7,1]$. With a negative skewness and parameters $w\in[3.1,6]$ the efficiency measures are always positive and exceed even 70%, as displayed in Graph 4.1. Finally, the ranking efficiency turns out to be a monotone decreasing function of the level of risk aversion (e.g. stronger effect for $m=1$ and weaker one for $m=4$).

Graph 4.1: NVG efficiency measure for monthly returns



4.2. Ranking efficiency for two stock market indices

In contrast to the benchmark data of Section 4.1, we consider now some stock market indices for which additionally the skewness and kurtosis can be estimated. Return observations stem from the following eight different Swiss Market (SMI) and Standard & Poors 500 (SP500) data sets:

SMI 3Y/1D: 758 historic daily closing prices over 3 years from 04.01.2010 to 28.12.2012
SMI 24Y/1D: 6030 historic daily closing prices over 12 years from 03.01.1989 to 28.12.2012
SMI 24Y/1M: 288 historic end of month prices over 24 years from Jan. 1989 to Dec. 2012
SP500 3Y/1D: 754 historic daily closing prices over 3 years from 04.01.2010 to 31.12.2012
SP500 24Y/1D: 6049 historic daily closing prices over 12 years from 03.01.1989 to 31.12.2012
SP500 24Y/1M: 288 historic end of month prices over 24 years from Jan. 1989 to Dec. 2012
SP500 63Y/1D: 15851 historic daily closing prices over 63 years from 03.01.1950 to 31.12.2012
SP500 63Y/1M: 756 historic end of month prices over 63 years from Jan. 1950 to Dec. 2012

These data sets are typical as they contain short to medium high volatile periods (recent 3 years), long term periods (24 years) as well as very long term periods (63 years). The SMI exists only for 24.5 years. Hence, the SMI cannot be compared with the SP500 for longer periods.

The observed sample logarithmic returns of stock-market indices are negatively skewed and have a much higher excess kurtosis than is allowed by a normal distribution, at least over shorter daily and even monthly periods. A simple test of rejection of the normal distribution is the Bera-Jarque (1987) statistic defined by

$$JB = n \cdot \left(\frac{\hat{S}}{6} + \frac{\hat{K}^2}{24} \right), \tag{4.5}$$

where n is the sample size, and $\hat{S}, \hat{K}$ are estimates of the skewness and (excess) kurtosis. This statistic is asymptotically χ_2^2 distributed, and has a critical value of 5.99 for a 95% confidence level. As seen from Table 4.2 below, the JB statistic is far beyond the critical value except for the monthly returns over 24 years (relatively small sample size of 288 observations). Therefore, the normal distribution is retained for comparison for the 3 monthly return data sets only. The VG and its NVG extension are fitted to the data following the moment method described in the Theorems 3.1 and 3.2. If the empirical counterparts of the domains of variation of the skewness and kurtosis are big enough, a unique solution is obtained, which is the case here.

To do so, the mean, variance, skewness and kurtosis, which are used in the moment method, must be estimated. We use the well-known *k-statistics* of Fisher (1928), which provide unbiased estimates of the cumulants as follows (assume $n > 3$):

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} r_i, \quad \hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(r_i - \hat{\mu})^2, \quad \hat{\kappa}_3 = \frac{n}{(n-1)(n-2)}\sum_{i=1}^{n}(r_i - \hat{\mu})^3, \quad \hat{S} = \hat{\kappa}_3 / \hat{\sigma}^3,$$
$$\hat{\kappa}_4 = \frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum_{i=1}^{n}(x_i - \hat{\mu})^4 - 3\cdot\frac{1}{(n-2)(n-3)}\left(\sum_{i=1}^{n}(x_i - \hat{\mu})^2\right)^2, \quad \hat{K} = \hat{\kappa}_4 / \hat{\sigma}^4, \tag{4.6}$$

where $r_i,\ i = 1,...,n,$ are the sample logarithmic returns (Table 4.2 lists the obtained values).

The *goodness-of-fit* (GoF) of the chosen estimation method is based on statistics, which measure the difference between the empirical distribution function $F_n(x)$ and the fitted distribution function $F(x)$. We use the Cramér-von Mises family of statistics defined by (e.g. D'Agostino and Stephens (1986), Cizek et al. (2005) and Burnecki *et al.* (2010))

$$T = n\cdot\int_{-\infty}^{\infty}\left[F_n(x) - F(x)\right]^2 w(x)dF(x), \tag{4.7}$$

where $w(x)$ is a suitable weighting function. If $w(x) = 1$ one obtains the W^2 *Cramér-von Mises* statistic (Cramér (1928), p.145-47, von Mises (1931), p.316-35). If $w(x) = 1/\left[F(x)\overline{F}(x)\right]$ one gets the A^2 *Anderson-Darling* statistic (Anderson and Darling (1952)). Consider the order statistics of the return data such that $r_1 \le r_2 \le ... \le r_n$ and let $\hat{F}(r_i),\ i = 1,...,n,$ be the fitted values of the distribution function. Then one has the formulas

$$W^2 = \frac{1}{12n} + \sum_{i=1}^{n}\left(\hat{F}(r_i) - \frac{2i-1}{n}\right)^2, \quad A^2 = -n - \sum_{i=1}^{n}\frac{2i-1}{n}\cdot\ln\left\{\hat{F}(r_i)\cdot\hat{\overline{F}}(r_{n-i+1})\right\}. \tag{4.8}$$

The fitted values $\hat{F}(r_i)$ are obtained numerically by integration of the expression (A4.21) for the VG, and by evaluation of (A4.28) for the NVG. The GoF statistic A^2 yields one of the most powerful test if the fitted distribution departs from the true distribution in the tails (e.g. D'Agostino and Stephens (1986)), and is recommended in this situation. Now, the observed sample logarithmic return data is negatively skewed and has a much higher kurtosis than is allowed by a normal distribution, which indicates that the fit in the tails matters and justifies the use of the GoF statistics (4.8). Needless to say, the moment method is only a starting point for improved GoF estimation methods, which include maximum likelihood estimation (MLE), minimum chi-square estimation, and miminization of the GoF statistics (4.8). However, such a more complex data analysis, undertaken by Hürlimann (2012) in another context, is beyond the scope of the present paper.

Fitting results are summarized and compared in the Table 4.1 below. Some comments are in order. Except for the SMI 24Y/1M data set, the NVG always provides the smallest GoF statistics. Sometimes the "best" fitted NVG is rather close to the VG (SP500 3Y/1D), close to VG (SMI 3Y/1D and SP500 24Y/1M), but also clearly departs from a VG (SMI 24Y/1D, SP500 24Y/1D, SP500 63Y/1D and SP500 63Y/1M). Even if the normal distribution is not rejected by the JB test, its fit is rather poor compared to the "best" NVG fit.

Since the NVG fits well the data for our purpose, no attempt has been made to compare the results with other test return distributions. Two classical competing analytically tractable choices are the normal inverse Gaussian and the generalized skew Student t (e.g. Aas and Haff (2006)), which have both been used successfully in Hürlimann (2009). To these models one can add another generalization of the Student t distribution by Hansen (1994) that has been studied by Jondeau and Rockinger (2003). Some comparisons about the feasible skewness and kurtosis limiting boundaries of these distributions are found in Appendix 3. Further interesting related choices include alternative skew Laplace versions and their extensions by Yu and Zhang (2005) and Wichitaksorn et al. (2012). The first authors construct the skew Laplace by combining two exponentials while the second author obtains it through a mixture of two scaled normal distributions. Another related family, which should be compared to the NVG, is the two-piece normal-Laplace family considered by Ardalan et al. (2012). Finally, it is important to mention that other methods and distributions are also able to capture the non-normality of return data, among others GARCH type models with a skewed t distribution, jump-diffusion models, generalized hyperbolic, stable and tempered stable distributions, as well as extreme value theory.

Table 4.1: Parameter estimates and GoF statistics for the NVG family

data set	parameter estimates						GoF statistics	
	s	υ	τ	ρ	α	β	A²	W²
SMI 3Y/1D	1	0.00079	0	0.85686	137.59	122.70	1.01399	0.16468
	0.93	0.00069	0.00372	0.64228	128.73	113.79	0.37267	0.04489
	0.915	**0.00067**	**0.00408**	**0.60214**	**126.83**	**111.88**	**0.35449**	**0.04407**
	0.90	0.00065	0.00441	0.56391	124.94	109.97	0.35799	0.04714
SP500 3Y/1D	1	0.00170	0	0.83167	122.17	101.43	0.40667	0.09112
	0.995	0.00168	0.00117	0.81544	121.65	100.90	0.34676	0.07623
	0.985	**0.00166**	**0.00202**	**0.78370**	**120.61**	**99.83**	**0.32735**	**0.07074**
	0.98	0.00164	0.00233	0.76818	120.10	99.30	0.34026	0.07326
SMI 24Y/1D	1	0.00075	0	0.42263	81.135	74.047	83.613	16.359
	0.83	0.00060	0.00663	0.20133	68.181	61.048	1.327	0.215
	0.82	**0.00059**	**0.00680**	**0.19186**	**67.422**	**60.285**	**1.232**	**0.210**
	0.81	0.00058	0.00697	0.18273	66.662	59.522	1.270	0.227
SP500 24Y/1D	1	0.00062	0	0.34901	74.756	69.580	94.937	18.683
	0.88	0.00054	0.00551	0.20961	66.186	60.997	3.252	0.421
	0.87	**0.00054**	**0.00572**	**0.20028**	**65.473**	**60.282**	**2.992**	**0.415**
	0.85	0.00052	0.00611	0.18254	64.046	58.853	3.153	0.519
SP500 63Y/1D	1	0.00066	0	0.11118	52.59512	44.65563	2185.79	354.64
	0.77	0.00051	0.00625	0.03981	42.16243	33.96995	25.44	1.83
	0.76	**0.00050**	**0.00637**	**0.03782**	**41.71793**	**33.50838**	23.34	**1.76**
	0.75	0.00050	0.00648	0.03592	41.27453	33.04714	22.05	1.81
	0.74	**0.00049**	**0.00659**	**0.03408**	**40.83230**	**32.58625**	**21.31**	1.95
	0.73	0.00049	0.00669	0.03232	40.39129	32.12571	21.32	2.18
SMI 24Y/1M	**1**	**0.03248**	**0**	**2.15686**	**62.759**	**35.041**	**0.32359**	**0.04054**
	0.99	0.03214	0.00677	2.08393	62.778	34.712	0.33634	0.04248
	0.98	0.03180	0.00955	2.01310	62.819	34.384	0.34940	0.04451
	0.95	0.03083	0.01499	1.81285	63.102	33.408	0.38922	0.05094
			normal distribution				2.60679	0.43223
SP500 24Y/1M	1	0.03468	0	2.48138	78.840	40.882	0.41544	0.07644
	0.93	0.03267	0.01595	1.96388	81.799	38.342	0.39113	0.07145
	0.92	**0.03243**	**0.01701**	**1.89885**	**82.520**	**37.991**	**0.39079**	**0.07141**
	0.91	0.03221	0.018	1.83596	83.348	37.642	0.39105	0.07150
			normal distribution				2.19806	0.37436
SP500 63Y/1M	1	0.01915	0	1.40715	50.036	33.985	1.38726	0.26121
	0.77	0.01478	0.02692	0.54708	45.817	26.233	0.25640	0.03680
	0.76	**0.01462**	**0.02742**	**0.52287**	**45.785**	**25.909**	**0.25430**	0.03586
	0.74	**0.01433**	**0.02838**	**0.47714**	**45.798**	**25.266**	0.25636	**0.03507**
	0.73	0.01419	0.02884	0.45558	45.848	24.947	0.26031	0.03518
			normal distribution				3.26036	0.51397

Let us now return to the main application, which is the evaluation of the efficiency measure (A1.8). Since the chosen estimation method is the moment method, the approximate ranking function $R_*^A(p)$ follows from (4.1) through direct insertion of the estimated paarameters. Moreover, to each solution (3.10)-(3.11) of the NVG moment problem, the corresponding test ranking function $R_*(p^{NVG})$ is evaluated using formula (4.4). In this way the ranking efficiency measure is obtained. The numerical results of our case study are summarized in Table 4.2. We note a systematic efficiency increase of the approximate ranking for the NVG over the Lévy and Markowitz (1979) benchmark. For each feasible value $s \in [|S| \cdot \sqrt{\frac{3}{2}/K}, 1]$ the efficiency increase is limited to a small range of variation. The maximum efficiency increase is here attained for the VG with $s = 1$ and the minimum for the NVG with $s = |S| \cdot \sqrt{\frac{3}{2}/K}$.

Table 4.2: NVG efficiency measures for SMI and SP500 data sets

data set	unbiased estimates				JB	min s	efficiency	
	μ	σ	S	K	statistic	|S|·√(1.5/K)	min	max
SMI 3Y/1D	0.00004	0.01011	-0.26118	3.54668	364	0.17027	92.83657	93.35457
SP500 3Y/1D	0.00031	0.01169	-0.42731	3.72928	383	0.27100	94.78351	95.04308
SMI 24Y/1D	0.00026	0.01189	-0.29736	7.15740	12568	0.13613	83.64588	86.15996
SP500 24Y/1D	0.00027	0.01160	-0.25716	8.63988	18552	0.10715	76.13769	81.19789
SP500 63Y/1D	0.00028	0.00980	-1.03074	27.69192	503713	0.23989	85.1725	87.14109
SMI 24Y/1M	0.00530	0.04800	-0.74866	1.77381	1.8148	0.68846	94.14608	94.27808
SP500 24Y/1M	0.00546	0.04340	-0.76442	1.61037	-5.5531	0.73776	95.35316	95.42011
SP500 63Y/1M	0.00586	0.04220	-0.65537	2.42167	102	0.51579	91.77448	92.20746

Appendix 1: Efficiency of portfolio selection ranking functions with CARA utility

Modern portfolio theory is mainly based on the classical mean-variance approach by Markowitz (1952) and subjective choice based on utility theory (von Neumann and Morgenstern (1947), Nash (1950)). To restrain the general problem by leaving out portfolio construction one focusses on portfolio selection only. Given is a finite set of portfolios, each with its own return distribution $p = p(x)$, and a rational investor with utility function $U(x)$, which is defined up to positive affine transformation. The *portfolio selection problem* consists to rank portfolios using the expected utility ranking function $R_U(p) = \int_{-\infty}^{\infty} U(x)p(x)dx$, or a function equivalent to it. Two ranking functions R_1 and R_2 are *equivalent*, written $R_1 \sim R_2$, if, and only if, there exists a monotone increasing function $h(x)$ such that $R_2(p) = h(R_1(p))$ for all p. Though the effects of higher moments on portfolio theory have been analyzed since many years (e.g. Samuelson (1970), Prakash et al. (2001), Jondeau and Rockinger (2006), Briec et al. (2007), Cvitanic et al. (2008), Jurczencko et al. (2008), Kleniati and Rustem (2009), Cesari and D'adda (2010)), so far not many simple results of general validity have been derived. Among the most important utility functions in use, one finds the constant absolute (CARA) and constant relative (CRRA) risk averse forms by Arrow (1965/71) and Pratt (1964)), the hyperbolic absolute risk averse form by Merton (1971), and the flexible three parameter (FTP) form considered by Conniffe (2007). Di Pierro and Mosevich (2011) attempt to clarify this issue for a rational risk-averse investor with a CARA utility function $U_{CARA}(x) = -\exp(-mx)$, also called *exponential utility*. For portfolio selection without risk-free asset, and assuming *finite moments*, these authors derive through a simple Taylor series expansion the *approximate* ranking equivalence such that

$$R_{U_{CARA}}(p) \sim R_*(p) = -\ln(-R_{U_{CARA}}(p))/m \approx R_*^A(p) = \mu - \frac{m\sigma^2}{2} + \frac{m^2\sigma^3 S}{6} - \frac{m^3\sigma^4 K}{720}, \quad \text{(A1.1)}$$

where the parameters μ, σ, S, K represent the mean, standard deviation, skewness and excess kurtosis of the portfolio return p, and the approximation error is of order $O(m^4\sigma^5)$. For Gaussian distributed return p^G equation (A1.1) reduces to the *exact* ranking function

$$R_*(p^G) = \mu - \frac{m\sigma^2}{2}, \quad \text{(A1.2)}$$

due to Lévy and Markowitz (1979). It is important to ask whether the approximate ranking function (A1.1) with *cubic* mean-variance-skewness-kurtosis trade-off should be preferred to the original ranking function with *linear* mean-variance trade-off (A1.2) or not.

To answer this question we examine the efficiency increase/decrease obtained using $R_*^A(p)$ instead of $R_*(p^G)$. For this, let S be an appropriate set of *test* return distributions, whose ranking functions $R_*(p) = -\log(-R_U(p))/m$ can be determined exactly or to an arbitrary level of accuracy for all $p \in S$. A *naive approach* to efficiency consists to measure the distance between two portfolio returns p_1 and p_2 through the *ranking distance*

$$D^*(p_1, p_2) = \left| R_*(p_1) - R_*(p_2) \right|. \quad \text{(A1.3)}$$

Now, let $R_*(p^G)$, resp. $R_*(p^S)$, be the ranking functions with Gaussian return, called *Gaussian ranking*, resp. test return $p^S \in S$, called *test ranking*, and let $R_*^A(p)$ be an *approximate ranking* valid for all returns p in a large set L of portfolio return distributions with finite moments. Then, a *ranking efficiency measure* of the approximate ranking versus the Gaussian ranking, given a test return $p^S \in S$, is described by the deviation of the distance measures $D^*(p^S, p)$ and $D^*(p^G, p)$, $\forall p \in L$, relative to the distance $D^*(p^G, p)$, in formula

$$E_S^*(p^G, p) = \frac{D^*(p^S, p) - D^*(p^G, p)}{D^*(p^G, p)} = \frac{\left|R_*(p^S) - R_*^A(p)\right|}{\left|R_*(p^G) - R_*^A(p)\right|} - 1, \forall p \in L,\ p^S \in S, \qquad \text{(A1.4)}$$

which quantifies the *efficiency increase* (if positive) respectively *decrease* (if negative) of the approximate ranking versus the Gaussian ranking. A similar approach, which is valid for any utility function $U(x)$, consists to replace (A1.3) by the *expected utility distance*

$$D^U(p_1, p_2) = \left|R_U(p_1) - R_U(p_2)\right|. \qquad \text{(A1.5)}$$

Now, similarly to the above, consider the quantities

$R_U(p^G)$: the expected utility with Gaussian return, or *Gaussian utility*

$R_U(p^S)$: the expected utility with test return, or *test utility*

$R_U^A(p)$: an approximation for the expected utility of returns $p \in L$, or *approximate utility*

Repeating the above explanation, one obtains the *expected utility efficiency measure*

$$E_S^U(p^G, p) = \frac{\left|R_U(p^S) - R_U^A(p)\right|}{\left|R_U(p^G) - R_U^A(p)\right|} - 1, \forall p \in L,\ p^S \in S. \qquad \text{(A1.6)}$$

Now, there is a well-known problem with the utility distance (A1.5). Since utility functions are only defined up to positive affine transformations, the naive distance (A1.5) is not appropriate (e.g. Lévy and Markowitz (1979), Hlawitschka (1994), etc.). A more meaningful approach must be considered (e.g. Kallbergaard and Ziemba (1979), Pulley (1983), etc.) and instead of expected utility the certainty equivalent return (CER) has been proposed as an alternative (e.g. Skoulakis (2012) for the CRRA utility or power utility used by Merton (1971)).

Let $CER(R_U(p)) = U^{-1}(R_U(p)) - 1$ be the CER associated to a utility function $U(x)$ with inverse $U^{-1}(x)$. Now, let us paraphrase the preceding construction. With the CER distance $D^{CER}(p_1, p_2) = \left|CER(R_U(p_1)) - CER(R_U(p_2))\right|$, and the CER quantities $CER(R_U(p^G))$, the *Gaussian CER*, $CER(R_U(p^S))$, the *test CER*, and $CER(R_U^A(p))$, $p \in L$, an *approximate CER*, one obtains, similarly to (A1.4) and (A1.6), the *CER efficiency measure*

$$E_S^{CER}(p^G, p) = \frac{\left|CER(R_U(p^S)) - CER(R_U^A(p)\right|}{\left|CER(R_U(p^G)) - CER(R_U^A(p)\right|} - 1, \forall p \in L,\ p^S \in S. \qquad \text{(A1.7)}$$

For a CARA utility function the following result holds.

Proposition A1.1: (*CER efficiency measure with CARA utility*). Suppose the expected utility of returns $p \in L$ is approximated by the formula $R_U^A(p) = U(R_*^A(p))$, where $U(x)$ is the CARA utility. Then the (naïve) ranking efficiency measure (A1.4) coincides with the CER efficiency measure (A1.7) such that

$$E_S^{CER}(p^G, p) = E_S^*(p^G, p) = \frac{\left|R_*(p^S) - R_*^A(p)\right|}{\left|R_*(p^G) - R_*^A(p)\right|} - 1, \forall p \in L,\ p^S \in S. \qquad \text{(A1.8)}$$

Proof: For CARA utility one has the equalities $R_U(p^S) = U\left(R_*(p^S)\right)$, $R_U(p^G) = U\left(R_*(p^G)\right)$. Since by assumption the equality $R_U^A(p) = U\left(R_*^A(p)\right)$ holds, the result follows by noting that

$$CER\left(R_U(p^S)\right) = R_*(p^S) - 1, \quad CER\left(R_U(p^G)\right) = R_*(p^G) - 1, \quad CER\left(R_U^A(p)\right) = R_*^A(p) - 1,$$

and inserting these equalities into (A1.7). ◊

Appendix 2: Skewness and kurtosis for the BG and NBG

Starting point are the BG scaled moment equations (3.3). We derive first two explicit parameterizations in terms of the scaled parameters. While the first one is function of the kurtosis only, the second one depends on both skewness and kurtosis. Their equivalence follows from the fact that the squared skewness is an explicit function of the kurtosis (later equation (A2.8)). It is convenient to use the following one-to-one transformation of the shape parameters

$$p = \gamma + \delta, \quad q^2 = \frac{\delta}{\gamma}, \quad p, q > 0, \quad \gamma = p \cdot \frac{1}{1+q^2}, \quad \delta = p \cdot \frac{q^2}{1+q^2}. \tag{A2.1}$$

Theorem A2.1: (*BG scale parameterizations*). The BG scale parameters (a, b) are *well-defined* if, and only if, one has $K \in [6p^{-1}, 6p^{-1}(1+q^{2\cdot \mathrm{sgn}(S)})]$ and the following equation holds:

$$c = p \cdot (1+q^2)^{-1} \cdot \left\{ \sqrt{\left(\frac{1+q\xi}{p}\right)^3} - q^2 \cdot \sqrt{\left(\frac{1-q^{-1}\xi}{p}\right)^3} \right\}, \quad \xi = \mathrm{sgn}(S) \cdot \sqrt{pd-1} \in [-q^{-1}, q], \tag{A2.2}$$

where by convention $\mathrm{sgn}(S=0) = 1$. The scale parameters are determined as follows:

<u>Scale parameterization (I)</u> (kurtosis dependence only)

$$a = \sqrt{\frac{1+q\xi}{p}}, \quad b = \sqrt{\frac{1-q^{-1}\xi}{p}}. \tag{A2.3}$$

<u>Scale parameterization (II)</u> (dependence on skewness and kurtosis)

$$\begin{aligned} a &= (pd)^{-1} \cdot \{(1+q\xi)c + (q-\xi) \cdot \sqrt{d-c^2}\}, \\ b &= (pd)^{-1} \cdot \{-(q-\xi)c + (1+q\xi) \cdot \sqrt{d-c^2}\}. \end{aligned} \tag{A2.4}$$

Proof: Comparing the scaled variance and kurtosis equations in (3.3) one obtains

$$a^2 = p^{-1} \cdot (1 \pm q\sqrt{pd-1}), \quad b^2 = p^{-1} \cdot (1 \pm q^{-1}\sqrt{pd-1}).$$

Taking into account the sign of the skewness according to the skewness equation in (3.3) one has with the defined quantity ξ in (A2.2) that

$$a^2 = p^{-1}(1+q\xi), \quad b^2 = p^{-1}(1-q^{-1}\xi). \tag{A2.5}$$

Both squares are non-negative and well-defined real numbers if, and only if, one has $\xi \in [-q^{-1}, q]$ and $\xi^2 = pd - 1 \geq 0$ (existence of square root in definition of ξ). Now, by definition of the parameter ξ, one has $\xi \in [0, q]$ if $S \geq 0$ and $\xi \in [-q^{-1}, 0]$ if $S < 0$. These conditions are equivalent with $d = p^{-1}(1+\xi^2) \in [p^{-1}, p^{-1}(1+q^{2\cdot \mathrm{sgn}(S)})]$ or $K \in [6p^{-1}, 6p^{-1}(1+q^{2\cdot \mathrm{sgn}(S)}]$, which settles the condition on the kurtosis. Inserting (A2.5) into the skewness equation $p \cdot (1+q^2)^{-1} \cdot (a^3 - q^2 \cdot b^3) = c$ yields further the condition (A2.2), and the first part of the result is shown. The parameterization (A2.3) is trivial in view of (A2.5). On the other hand, taking into account (A2.5), the skewness equation can be rewritten as

$$(q-\xi)qb = (1+q\xi)a - (1+q^2)c.$$

Squaring, multiplying by p and using the variance relationship $pq^2b^2 = 1+q^2 - pa^2$ one obtains the quadratic equation

$$p\{(1+q\xi)^2 + (q-\xi)^2\}a^2 - 2pc(1+q^2)(1+q\xi)a - (1+q^2)(q-\xi)^2 + (1+q^2)^2 pc^2 = 0,$$

or equivalently

$$p(1+\xi^2)a^2 - 2pc(1+q\xi)a - (q-\xi)^2 + (1+q^2)pc^2 = 0.$$

The reduced discriminant is equal to

$$\Delta = p(1+\xi^2)(q-\xi)^2 - p^2c^2\{(1+q^2)(1+\xi^2) - (1+q\xi)^2\} = p^2(q-\xi)^2(d-c^2),$$

where, in the first term, use is made of the relation $1+\xi^2 = pd$. Since $\Delta \geq 0$ by the inequality of Theorem A2.2 below, the solution of this quadratic equation yields the expression for a (and similarly for b) in (A2.4), where a priori both signs ± for the square root term are possible. A final calculation shows that the + sign is adequate such that (A2.4) solves (3.3). ◊

Theorem A2.2: (*BG inequality between skewness and kurtosis*). The inequality $S^2 \leq \frac{2}{3}K$ is sharp and attained at the pairs $(S,K) = (2\varepsilon\sqrt{p^{-1}(1+q^{2\varepsilon})}, 6p^{-1}(1+q^{2\varepsilon}))$, $\varepsilon = \operatorname{sgn}(S)$, that is for the limiting left- and right-tail gamma distributions $\ell G(\upsilon^-,\rho^-,\alpha^-)$ and $rG(\upsilon^+,\rho^+,\alpha^+)$ with $\upsilon^{\pm} = \mu \pm \sigma\cdot\sqrt{p(1+q^{\pm 2})^{-1}}$, $\rho^{\pm} = p(1+q^{\pm 2})^{-1}$, $\alpha^{\pm} = \sigma^{-1}\cdot\sqrt{p(1+q^{\pm 2})^{-1}}$, and $\pm$ stands for ε.

Proof: Recall that c satisfies the equation in (A2.2) and that $pd = 1+\xi^2$. To derive the inequality $S^2 \leq \frac{2}{3}K$, or equivalently $c^2 \leq d$, one must therefore show that

$$\left[\sqrt{\left(\frac{1+q\xi}{1+q^2}\right)^3} - q^2\cdot\sqrt{\left(\frac{1-q^{-1}\xi}{1+q^2}\right)^3}\right]^2 \leq \frac{1+\xi^2}{1+q^2}, \quad \forall\xi\in[-q^{-1},q], \forall q>0. \qquad \text{(A2.6)}$$

For this it suffices to discuss the analytic properties of the auxiliary function

$$f(x) = \frac{1+x^2}{1+q^2} - \left[\left(\frac{1+qx}{1+q^2}\right)^{3/2} - q^2\left(\frac{1-q^{-1}x}{1+q^2}\right)^{3/2}\right]^2, \quad x\in[-q^{-1},q].$$

Set $g(x) = (1+q^2)^3 f(x) = (1+q^2)^2(1+x^2) - \left[(1+qx)^{3/2} - q^2(1-q^{-1}x)^{3/2}\right]^2$ to see that

$$g(x) = 2q^2 + 3q(q^2-1)x + (1-4q^2+q^4)x^2 - q(q^2-1)x^3 + 2q(1+qx)(q-x)\sqrt{(1+qx)(1-q^{-1}x)}.$$

Since $g(\pm q^{\pm}) = 0$ and the square-root term is divisible by the product of linear factors $(1+qx)(q-x)$, the polynomial expression of third degree must be divisible by these linear factors. A calculation shows that $g(x) = (1+qx)(q-x)h(x)$ with

$$h(x) = (q^2-1)x + 2q + 2q\sqrt{(1+qx)(1-q^{-1}x)}.$$

Now, if $x\in[-q^{-1},q]$ we have $x \geq -q^{-1}, -x \geq -q$, and we see that

$$h(x) \geq -q^2q^{-1} - q + 2q + 2q\sqrt{(1+qx)(1-q^{-1}x)} \geq 2q\sqrt{(1+qx)(1-q^{-1}x)} \geq 0.$$

Therefore $f(x) \geq 0$ on $[-q^{-1}, q]$, which is (A2.6). Since $f(\pm q^{\pm 1}) = 0$ the inequality is sharp and attained exactly when $\xi = \pm q^{\pm 1}$. With (A2.2) this yields $d = p^{-1}(1+q^{\pm 2}), c = \pm p^{-1}(1+q^{\pm 2})$, or $K = 6p^{-1}(1+q^{\pm 2}), S = \pm 2\sqrt{p^{-1}(1+q^{\pm 2})}$. Moreover, with (A2.3) the case $\xi = -q^{-1}$ implies $a = 0$, or $\alpha \to \infty$, and $b = \sqrt{p^{-1}(1+q^{-2})}$. The case $\xi = q$ implies $b = 0$, or $\beta \to \infty$, and $a = \sqrt{p^{-1}(1+q^2)}$. These conditions characterize the limiting left- and right-tail gamma distributions. The values of the parameters follow by inserting the scale parameters (a, b) into the moment equations (3.3). ◊

Remarks A2.1: Theorem A2.2 includes the important VG with parameters $q = 1,\ p = 2\rho$. In this special case the inequality between skewness and kurtosis has also been derived in Ghysels and Wang (2011), Proposition 2.4. It is important to ask whether the whole domain of variation between skewness and kurtosis can be attained for some of its members. A positive answer has been provided in Theorem 3.1.

Corollary A2.1: (*Symmetric BG scale parameterizations*). The scale parameters of the symmetric generalized Laplace $sGL(\mu, \rho, a) = BG(\nu = \mu, \gamma = \rho, a, \delta = \rho, b = a)$, are well-defined if, and only if, one has $(S, K) = (0, 6p^{-1})$, and $(\rho, a) = (\frac{1}{2}p, \sqrt{p^{-1}})$.

Proof: In case $b = a$ one has by (A2.3) that $(q + q^{-1})\xi = 0$, hence $\xi = 0$, and thus $b = a = \sqrt{p^{-1}}$. By symmetry $S = 0$ and $K = 6p^{-1}$ follows because $pd = 1$ in (A2.2). ◊

Since ξ depends on the kurtosis, one notes that (A2.2) is an *exact* relationship between skewness and kurtosis. Explicitly, squaring (A2.2) and multiplying with $p(1+q^2)^2$ one gets

$$\begin{aligned} p(1+q^2)^2c^2 &= (\sqrt{(1+q\xi)^3} - \sqrt{q(q-\xi^2)^3})^2 \\ &= (1+q^4) - 3q(q^2-1)\xi + 6q^2\xi^2 + q(q^2-1)\xi^3 - 2\sqrt{q(1+q\xi)^3(q-\xi^2)^3} \\ &= (1+q^2)^2(1+\xi^2) - g(\xi), \end{aligned}$$

where for the last equality the auxiliary function $g(x)$ from the proof of Theorem A2.2 has been used. A further use of the factorization $g(x) = (1+qx)(q-x)h(x)$ yields the equation

$$\begin{aligned} & p(1+q^2)^2c^2 \\ &= (1+q^2)^2(1+\xi^2) - (1+q\xi)(q-\xi)((q^2-1)\xi + 2q) - 2\sqrt{q(1+q\xi)^3(q-\xi^2)^3}, \\ &\xi = \operatorname{sgn}(S)\cdot\sqrt{6^{-1}(pK-6)}, \quad pK \in [6, 6(1+q^{2\cdot\operatorname{sgn}(S)})]. \end{aligned} \tag{A2.7}$$

The VG special case $q = 1,\ p = 2\rho$ simplifies considerably. From (A2.7) and the relationship $2\rho d = 1 + \xi^2$ one gets

$$c^2 = \frac{1}{2\rho}(3\rho d - 1 - \sqrt{2(1-\rho d)^3}). \tag{A2.8}$$

Theorem A2.3: (*VG equation of skewness/kurtosis and extremal values*). The squared skewness is an increasing concave function of the kurtosis, which satisfies the sharp inequalities

$$0 \leq S^2 = \rho^{-1} \cdot \left\{ \rho K - 2 - \sqrt{\left(\frac{6-\rho K}{3}\right)^3} \right\} \leq 4\rho^{-1}, \quad \rho K \in [3, 6]. \tag{A2.9}$$

The minimum squared skewness is attained for a symmetric generalized Laplace $(S, K) = (0, 3\rho^{-1})$, and the maximum for the left- and right-tail gamma distributions with $(S, K) = (\pm 2\sqrt{\rho^{-1}}, 6\rho^{-1})$.

Proof: Descaling (A2.8) with the transformation (3.2), one obtains the equation of skewness and kurtosis in (A2.9). On the other hand, the auxiliary function $h(x) = 3\rho x - 1 - \sqrt{2(1-\rho x)^3}$ associated to (A2.8) is increasing concave for $\rho x \in \left[\frac{1}{2}, 1\right]$. Therefore, the extremal values of (A2.7) are attained at $h(\frac{1}{2}\rho^{-1}) = 0$ and $h(\rho^{-1}) = 2$. This implies the inequalities in (A2.9). The fact that the pairs $(S,K) = (0, 3\rho^{-1})$, $(S,K) = (\pm 2\sqrt{\rho^{-1}}, 6\rho^{-1})$ belong to the symmetric generalized Laplace respectively to the left- and right-tail gamma follows from Corollary A2.1 respectively Theorem A2.2. ◊

Starting point of the analysis for the NBG are the scaled moment equations (3.9). Since this system has the same form as the BG system (3.3), a great part of the NBG analysis directly follows from the BG case. Theorem A2.1 generalizes as follows to the NBG family.

Theorem A2.4: (*NBG scale parameterizations*). For each $s \in (0,1]$ the scale parameters $(\overline{a}, \overline{b})$ of the normal bilateral gamma $NBG(\nu, \tau = \sigma\sqrt{1-s^2}, \gamma, \overline{a}s, \delta, \overline{b}s)$ are *well-defined* if, and only if, one has $K \in [6s^4 p^{-1}, 6s^4 p^{-1}(1+q^{2\cdot \mathrm{sgn}(S)})]$ and the following equation holds:

$$\overline{c} = p\cdot(1+q^2)^{-1}\cdot\left\{\sqrt{\left(\frac{1+q\overline{\xi}}{p}\right)^3} - q^2\cdot\sqrt{\left(\frac{1-q^{-1}\overline{\xi}}{p}\right)^3}\right\}, \quad \overline{\xi} = \mathrm{sgn}(S)\cdot\sqrt{p\overline{d}-1} \in [-q^{-1}, q], \quad \text{(A2.10)}$$

The scale parameters are determined as follows:

Scale parameterization (I) (kurtosis dependence only)

$$\overline{a} = \sqrt{\frac{1+q\overline{\xi}}{p}}, \quad \overline{b} = \sqrt{\frac{1-q^{-1}\overline{\xi}}{p}}. \quad \text{(A2.11)}$$

Scale parameterization (II) (dependence on skewness and kurtosis)

$$\begin{aligned}\overline{a} &= (p\overline{d})^{-1}\cdot\{(1+q\overline{\xi})\overline{c} + (q-\overline{\xi})\cdot\sqrt{\overline{d}-\overline{c}^2}\},\\ \overline{b} &= (p\overline{d})^{-1}\cdot\{-(q-\overline{\xi})\overline{c} + (1+q\overline{\xi})\cdot\sqrt{\overline{d}-\overline{c}^2}\}.\end{aligned} \quad \text{(A2.12)}$$

Proof: It suffices to rewrite Theorem A2.1 in terms of the iterated scale parameters. The kurtosis condition follows from the fact $\overline{d} \in [p^{-1}, p^{-1}(1+q^{2\cdot \mathrm{sgn}(S)})]$ in the proof of Theorem A2.1. ◊

Theorem A2.5: (*NBG inequality between skewness and kurtosis*). The inequality $S^2 \le \frac{2}{3}Ks^2$, $s \in (0,1]$ is sharp and attained at $(S,K) = (2\varepsilon\cdot s^3\sqrt{p^{-1}(1+q^{2\varepsilon})}, 6s^4 p^{-1}(1+q^{2\varepsilon}))$, $\varepsilon = \mathrm{sgn}(S)$, that is for the limiting left- and right-tail normal gamma distributions $\ell NG(\upsilon^-, \tau = \sigma\sqrt{1-s^2}, \rho^-, \alpha^-)$, $rNG(\upsilon^+, \tau = \sigma\sqrt{1-s^2}, \rho^+, \alpha^+)$, $\upsilon^{\pm} = \mu \pm s\sigma\cdot\sqrt{p(1+q^{\pm 2})^{-1}}$, $\rho^{\pm} = p(1+q^{\pm 2})^{-1}$, $\alpha^{\pm} = (s\sigma)^{-1}\cdot\sqrt{p(1+q^{\pm 2})^{-1}}$, and $\pm$ stands for the sign of skewness.

Proof: The inequality $\overline{d} \ge \overline{c}^2$ is the counterpart of the inequality $d \ge c^2$ in the proof of Theorem A2.2. The inequality $S^2 \le \frac{2}{3}Ks^2$ follows by de-scaling the parameters using that $\overline{c} = cs^{-3} = \frac{1}{2}Ss^{-3}$, $\overline{d} = ds^{-4} = \frac{1}{6}Ks^{-4}$. The inequality is sharp and attained exactly when $\overline{\xi} = \pm q^{\pm 1}$, i.e. $\overline{d} = p^{-1}(1+q^{\pm 2})$, $\overline{c} = \pm p^{-1}(1+q^{\pm 2})$, or $K = 6s^4 p^{-1}(1+q^{\pm 2})$, $S = \pm 2s^3\sqrt{p^{-1}(1+q^{\pm 2})}$. The case $\overline{\xi} = -q^{-1}$ implies $\overline{a} = 0$, or $\alpha \to \infty$, and $\overline{b} = \sqrt{p^{-1}(1+q^{-2})}$. The case $\overline{\xi} = q$ implies $\overline{b} = 0$, or $\beta \to \infty$, and $\overline{a} = \sqrt{p^{-1}(1+q^2)}$. These conditions characterize the limiting left- and right-tail normal gamma distributions. The values of the parameters follow by inserting the scale parameters $(\overline{a}, \overline{b})$ into the moment equations (3.9) and de-scaling the parameters. ◊

Remark A2.2: The generic NBG inequality $S^2 \le \frac{2}{3}Ks^2,\ s \in (0,1]$, is more restricted than the BG inequality stated in Theorem A2.2. However, in case the parameter s is sufficiently high, the NBG is enough flexible for modelling purposes, as demonstrated in Section 4.2.

Similarly to (A2.7) the equation (A2.10) can be rewritten as

$$\begin{aligned} &p(1+q^2)^2\bar{c}^2 = (1+q^2)^2(1+\bar{\xi}^2) - (1+q\bar{\xi})(q-\bar{\xi})((q^2-1)\bar{\xi}+2q) - 2\sqrt{q(1+q\bar{\xi})^3(q-\bar{\xi}^2)^3}, \\ &\bar{\xi} = \mathrm{sgn}(S)\cdot\sqrt{6^{-1}(ps^{-4}\gamma_2 - 6)}, \quad ps^{-4}K \in [6, 6(1+q^{2\cdot\mathrm{sgn}(S)})]. \end{aligned} \tag{A2.13}$$

The NVG special case $q=1,\ p=2\rho$ simplifies. From (A2.13) and the relationship $2\rho\bar{d} = 1+\bar{\xi}^2$ one gets

$$\bar{c}^2 = \frac{1}{2\rho}(3\rho\bar{d} - 1 - \sqrt{2(1-\rho\bar{d})^3}). \tag{A2.14}$$

Theorem A2.6: (*NVG equation of skewness/kurtosis and extremal values*). The squared skewness is an increasing concave function of the kurtosis, which satisfies the sharp inequalities

$$0 \le S^2 = \rho^{-1}\cdot\left\{(\rho K - 2s^4)s^2 - \sqrt{\left(\frac{6s^4-\rho K}{3}\right)^3}\right\} \le 4\rho^{-1}s^6, \quad \rho K \in \left[3s^4, 6s^4\right],\ s \in (0,1]. \tag{A2.15}$$

The minimum squared skewness is attained for a symmetric normal generalized Laplace $sNGL(\nu=\mu, \tau=\sqrt{1-s^2}\sigma, \alpha=(s\sigma)^{-1}\sqrt{2\rho}, \beta=\alpha)$ with $(S,K)=(0, 3\rho^{-1}s^4)$, and the maximum for the left- and right-tail normal gamma with $(S,K)=(\pm 2s^3\sqrt{\rho^{-1}}, 6\rho^{-1}s^4)$.

Proof: The equation (A2.15) follows from (A2.14) by de-scaling with $\bar{c} = \frac{1}{2}Ss^{-3}$, $\bar{d} = \frac{1}{6}Ks^{-4}$. The extremal values of $\bar{c}^2$ are 0 and ρ^{-1} and are attained at $\bar{d} = \frac{1}{2}\rho^{-1}$ and $\bar{d} = \rho^{-1}$. In case $\bar{d} = \frac{1}{2}\rho^{-1}$ one has $(S,K)=(0,3\rho^{-1}s^4),\ \bar{\xi}=0,\ a=b=s\sqrt{\frac{1}{2}\rho^{-1}}$, which is the normal symmetric generalized Laplace. The case $\bar{d}=\rho^{-1}$ corresponds to the pairs $(S,K)=(\pm 2s^3\sqrt{\rho^{-1}}, 6\rho^{-1}s^4)$, which are the left- and right-tail normal gamma according to Theorem A2.5. ◊

Appendix 3: Comparison of some skewness and kurtosis boundaries

<u>Comparison with the domain of maximum size</u>

It is instructive to compare the BG inequality of Theorem A2.2 with the general inequality between skewness and kurtosis for arbitrary distributions on $(-\infty,\infty)$, namely

$$K \ge S^2 - 2 \quad \text{or equivalently} \quad c^2 \le \frac{1}{2}(1+3d), \tag{A3.1}$$

which is sharp and attained at a biatomic random variable with support $\{\omega, \bar{\omega} = -\omega^{-1}\}$, where $\omega = \frac{1}{2}(S - \sqrt{4+S^2})$ (Pearson (1916), Wilkins (1944), Guiard (1980), Hürlimann (2008b), Theorem I.4.1). A family of distributions, which is able to model any admissible pair (S,K), is the *Johnson system* introduced in Johnson (1949) (see also Johnson et al. (1994), George (2007) among others). Note that for distributions with a finite range $[A,B], -\infty < A < B < \infty$, the inequality (A3.1) extends to a two-sided inequality (furthermore information is found in Hürlimann (2008b), Chap.I.4). Clearly, the domain of variation of skewness and kurtosis for the BG, and a fortiori NBG, is much more restricted than the domain of maximum size prescribed by the inequality (A3.1). This follows because trivially $c^2 \le d \le \frac{1}{2}(1+3d)$ is always satisfied.

Comparison with the normal inverse Gaussian and the generalized skew t

The domain of variation between skewness and kurtosis is larger for the BG than for the *normal inverse Gaussian* (NIG). Indeed, consider the NIG random variable $X = \mu + \delta \cdot Z \sim NIG(\mu, \delta, \alpha, \beta)$, with $Z \sim SNIG(\alpha, \beta)$ a standard NIG random variable such that $\mu = 0, \delta = 1$. The cgf of the latter equals

$$C_Z(t) = \sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + t)^2} \ .$$

Setting $\zeta = \beta / \alpha$ one expresses the squared skewness and kurtosis as

$$S^2 = \frac{9\zeta^2}{\alpha\sqrt{1-\zeta^2}}, \quad K = \frac{3(1+4\zeta^2)}{\alpha\sqrt{1-\zeta^2}}, \text{ hence } 0 \le \zeta^2 = \frac{S^2}{3K - 4S^2} \le 1,$$

which shows that the NIG domain $S^2 \le \frac{3}{5}K$ is contained in the BG domain $S^2 \le \frac{2}{3}K$. The corresponding result for the VG special case is found in Ghysels and Wang (2011), p.8. These authors also show that the NIG domain contains the domain of the *generalized skew t distribution* (GST) (applications of the GST are found in Frecka and Hopwood (1983), Theodossiu (1998), Aas and Haff (2006), Hürlimann (2009), etc.).

Comparison with Hansen's generalized t

Hansen (1994) considers another generalization of the Student t distribution, simply called *generalized t* (GT) by Jondeau and Rockinger (2003). By the Theorems A2.2 and A2.5 the boundaries of maximum skewness by given kurtosis for the BG and NBG are delimited by the two curves $S = \pm\sqrt{\frac{2}{3}Ks}, \ s \in (0,1]$. This domain is in particular maximum for the BG and VG, both realized when $s = 1$. In this situation, let $S_{BG}^2(K) = \frac{2}{3}K$ denote the maximum squared skewness as a function of K. The GT skewness and kurtosis boundary has been determined in Jondeau and Rockinger (2003), Section 2.2, Fig. 5 (note that the excess kurtosis is obtained by subtracting the constant 3 from the expression (3) in Section 2.1). Let $S_{GT}^2(K)$ denote the corresponding maximum squared skewness. The GT domain is contained in the BG (and VG) domain for kurtosis higher than some relatively moderate value. One has

$$S_{BG}^2(K) \le S_{GT}^2(K), \quad \forall\ K \le K_0 = 2.774, \quad S_{BG}^2(K) > S_{GT}^2(K), \quad \forall\ K > K_0. \qquad \text{(A3.3)}$$

Finally, the NBG kurtosis is bounded from below by zero, i.e. the NBG does not allow tails to be thinner than those of the normal (limiting case $\alpha \to \infty, \beta \to \infty$ of the NL in the Notes 2.1). This property is shared with the GT, as observed by Jondeau and Rockinger (2003), Section 2.2.

Appendix 4: Special function representations of densities and distributions

It suffices to restrict the attention to the BG with vanishing location $\upsilon = 0$. The BG pdf, denoted by $f(x) = f(x; \gamma, \alpha, \delta, \beta)$, is the convolution $f(x) = (f_1 * f_2)(x)$ of the two gamma pdf's defined by

$$f_1(x) = \Gamma(\gamma)^{-1}\alpha^\gamma x^{\gamma-1}e^{-\alpha x} \cdot 1\{x \ge 0\}, \quad f_2(x) = \Gamma(\delta)^{-1}\beta^\delta |x|^{\delta-1}e^{-\beta|x|} \cdot 1\{x \le 0\}. \qquad \text{(A4.1)}$$

A formula from Oldham et al. (2009) will be referred to as a formula from Atlas (2009).

Generalized gamma or Tricomi function representation

The first "generalized gamma function" representation seems new. It is equivalent to the representation (A4.6) below in terms of the confluent hyper-geometric function of the 2nd kind.

Theorem A4.1: (*Generalized gamma function representation*). The probability density function of the bilateral gamma $BG(\upsilon = 0, \gamma, \alpha, \delta, \beta)$ is given by

$$\begin{aligned} f(x) &= \Gamma(\gamma)^{-1}\Gamma(\delta)^{-1}\left(\tfrac{\beta}{\alpha+\beta}\right)^\delta \alpha^\gamma x^{\gamma-1}e^{-\alpha x} \cdot \Gamma(\delta, \gamma, (\alpha+\beta)x), \quad x > 0, \\ f(x) &= \Gamma(\gamma)^{-1}\Gamma(\delta)^{-1}\left(\tfrac{\alpha}{\alpha+\beta}\right)^\gamma \beta^\delta |x|^{\delta-1}e^{-\beta|x|} \cdot \Gamma(\gamma, \delta, (\alpha+\beta)|x|), \quad x < 0, \end{aligned} \qquad \text{(A4.2)}$$

with the *generalized gamma function*

$$\Gamma(a,b,x)=\int_0^\infty t^{a-1}(1+x^{-1}t)^{b-1}e^{-t}dt. \qquad \text{(A4.3)}$$

Proof: Using the symmetry relation $f(x;\gamma,\alpha,\delta,\beta)=f(-x;\delta,\beta,\gamma,\alpha)$ it suffices to consider the case $x\in(0,\infty)$. Through elementary integration (change of variables $y=-tx$) one obtains

$$f(x)=(f_1*f_2)(x)=\int_{-\infty}^0 f_1(x-y)f_2(y)dy=\Gamma(\gamma)^{-1}\Gamma(\delta)^{-1}\alpha^\gamma\beta^\delta e^{-\alpha x}\cdot I(x),$$
$$I(x)=\int_{-\infty}^0 (x-y)^{\gamma-1}(-y)^{\delta-1}e^{(\alpha+\beta)y}dy=\int_0^\infty x^{\gamma+\delta-1}(1+t)^{\gamma-1}t^{\delta-1}e^{-(\alpha+\beta)xt}dt,$$

The transformation $t=c(x)^{-1}u$ with $c(x)=(\alpha+\beta)x$ yields further

$$I(x)=x^{\gamma+\delta-1}c(x)^{-\delta}\cdot\int_0^\infty(1+c(x)^{-1}u)^{\gamma-1}u^{\delta-1}e^{-u}du=x^{\gamma-1}(\alpha+\beta)^{-\delta}\cdot\Gamma(\delta,\gamma,c(x)).$$

Insert into the first integral expression for $f(x)$ to get (A4.2). ◊

Remarks A4.1: In virtue of the limiting property $\lim_{x\to\infty}\Gamma(a,b,x)=\int_0^\infty t^{a-1}e^{-t}dt=\Gamma(a)$ the naming of the integral (A4.3) is justified. Furthermore, one has also trivially $\Gamma(a,1,x)=\Gamma(a)$. Another justification arises from the fact that when $\alpha\to\infty$ or $\beta\to\infty$ the pdf converges to a left- and right-tail gamma pdf respectively, as should be. Moreover, a close look at the *confluent hyper-geometric function of the 2nd kind*, introduced by Tricomi (1947) and also called Tricomi function, shows the relationship

$$\Gamma(a,b,x)=\Gamma(a)x^aU(a,a+b,x), \qquad \text{(A4.4)}$$

where the *Tricomi function* is defined by (e.g. Atlas (2009), 48:3:6 and 48:3.7)

$$U(a,b,x)=\Gamma(a)^{-1}\cdot\int_0^\infty t^{a-1}(1+t)^{b-a-1}e^{-xt}dt=\Gamma(a)^{-1}\cdot\int_0^1 t^{a-1}(1-t)^{-b}e^{-xt(1-t)^{-1}}dt. \qquad \text{(A4.5)}$$

The generalized gamma function is a transformed Tricomi function and (A4.2) rewrites as

$$\begin{aligned} f(x)&=\Gamma(\gamma)^{-1}\alpha^\gamma x^{\gamma-1}e^{-\alpha x}(\beta x)^\delta\cdot U(\delta,\gamma+\delta,(\alpha+\beta)x), \quad x>0,\\ f(x)&=\Gamma(\delta)^{-1}\beta^\delta|x|^{\delta-1}e^{-\beta|x|}(\alpha|x|)^\gamma\cdot U(\gamma,\gamma+\delta,(\alpha+\beta)|x|), \quad x<0.\end{aligned} \qquad \text{(A4.6)}$$

Further alternative special function representations of the BG density are available.

<u>Kummer function representation</u>

Provided b is not an integer, the Tricomi function (A4.5) can be represented as a weighted sum of two Kummer functions. Otherwise, the situation can be handled using a complicated formula (e.g. Atlas (2009), 48:3:1 and 48:3:3)

Case 1: $a,b>0,\ a+b\notin N$

$$U(a,a+b,x)=\frac{\Gamma(1-a-b)}{\Gamma(1-b)}M(a,a+b,x)+\frac{\Gamma(a+b-1)}{\Gamma(a)}x^{1-a-b}M(1-b,2-a-b,x) \qquad \text{(A4.7)}$$

where the *Kummer function* is defined by the convergent power series expansion

$$M(a,b,x)=\sum_{k=0}^\infty\frac{(a)_k}{(b)_k k!}\cdot x^k, \quad (y)_k=\frac{\Gamma(y+k)}{\Gamma(y)} \quad \text{(Pochhammer symbol).} \qquad \text{(A4.8)}$$

Case 2: $a,b>0,\ a+b=n\in N$

$$U(a,n,x)=\frac{(n-2)!}{\Gamma(a)}x^{1-n}\cdot\sum_{k=0}^{n-2}\frac{(1+a-n)_k}{(2-n)_k\,k!}\cdot x^k$$
$$+\frac{(-1)^n}{(n-1)!\Gamma(1+a-n)}\cdot\sum_{j=0}^{\infty}[\psi(j+a)-\psi(j+1)-\psi(j+n)+\ln(x)]\frac{(a)_{\mathrm{j}}}{(n)_{\mathrm{j}}\,j!}\cdot x^j, \tag{A4.9}$$

with the digamma function $\psi(\cdot)$. Simplification occurs if $n=1,2$.

With this the Tricomi representation (A4.6) translates to the Kummer representation (for simplicity only Case 1 with $\gamma+\delta\notin N$ for $x>0$ is reproduced here):

$$f(x)=\Gamma(\gamma)^{-1}\alpha^{\gamma}x^{\gamma-1}e^{-\alpha x}\cdot(\beta x)^{\delta}\cdot$$
$$\left\{\begin{array}{l}\dfrac{\Gamma(1-\gamma-\delta)}{\Gamma(1-\gamma)}M(\delta,\gamma+\delta,(\alpha+\beta)x)\\ +\dfrac{\Gamma(\gamma+\delta-1)}{\Gamma(\delta)}[(\alpha+\beta)x]^{1-\gamma-\delta}M(1-\gamma,2-\gamma-\delta,(\alpha+\beta)x)\end{array}\right\},\quad x>0. \tag{A4.10}$$

In view of the series (A4.8) the Kummer representation is suitable for numerical evaluation.

<u>Whittaker function representation</u>

Whittaker has introduced the normalized versions $M_{\nu,\mu}(x)$ and $W_{\nu,\mu}(x)$ of the Kummer and Tricomi functions. They depend upon the following one-to-one parameter transformation:

$$\nu=\tfrac{1}{2}b-a,\qquad \mu=\tfrac{1}{2}(b-1),\quad a=\tfrac{1}{2}+\mu-\nu,\quad b=1+2\mu. \tag{A4.11}$$

The Whittaker functions relate to the Kummer and Tricomi functions through the relationships (e.g. Atlas (2009), 48:13:2)

$$U(a,b,x)=x^{-\frac{1}{2}b}e^{\frac{1}{2}x}W_{\nu,\mu}(x),\quad M(a,b,x)=x^{-\frac{1}{2}b}e^{\frac{1}{2}x}M_{\nu,\mu}(x). \tag{A4.12}$$

Some relationships are more symmetrical in Whittaker's notation (e.g. Atlas (2009), 48:13:3 and 48:13:4)

$$W_{\nu,-\mu}(x)=W_{\nu,\mu}(x),\quad M_{\nu,\mu}(-x)=(-1)^{\mu+\frac{1}{2}}M_{-\nu,\mu}(x). \tag{A4.13}$$

Inserting (A4.12) into (A4.6) one obtains the Whittaker representation (here for $x>0$ only)

$$f(x)=\Gamma(\gamma)^{-1}\frac{\alpha^{\gamma}\beta^{\delta}}{(\alpha+\beta)^{\frac{1}{2}(\gamma+\delta)}}x^{\frac{1}{2}(\gamma+\delta)-1}e^{-\frac{1}{2}(\alpha-\beta)x}\cdot W_{\frac{1}{2}(\gamma-\delta),\frac{1}{2}(\gamma+\delta)-\frac{1}{2}}((\alpha+\beta)x),\quad x>0. \tag{A4.14}$$

This representation is displayed in Kücher and Tappe (2008a/b) as formula (3.4) respectively (4.4). Similarly to (A4.7) the Whittaker function $W_{\nu,\mu}(x)$ can be expressed as weighted sum of two Whittaker functions $M_{\nu,\mu}(x)$ as follows (e.g. Gradshteyn and Ryzhik (2000), p.1014):

$$W_{\nu,\mu}(x)=\frac{\Gamma(-2\mu)}{\Gamma(\frac{1}{2}-\nu-\mu)}M_{\nu,\mu}(x)+\frac{\Gamma(2\mu)}{\Gamma(\frac{1}{2}-\nu+\mu)}M_{\nu,-\mu}(x). \tag{A4.15}$$

Since (A4.15) can be expressed via (A4.12) as function of the convergent series (A4.8), the expression (A4.14) can be evaluated numerically. This is mentioned in Küchler and Tappe (2008a/b) without the required restriction to non-integer values of $\gamma+\delta$, however.

Remark A4.2: It is also possible to express the BG density in terms of Dirichlet B-spline functions as advocated in Kaishev (2010).

Examples A4.1: BG sub-families and the VG density

There exist a number of important special cases under which the presented special function representations simplify considerably and are mathematically much more tractable.

Bilateral exponential gamma and bilateral gamma exponential

The shape parameters are $(\gamma = 1, \delta > 0)$ and $(\gamma > 0, \delta = 1)$. We need the special cases of the Tricomi functions (Atlas (2009), 48:4:2 and 48:4:6):

$$U(a, a+1, x) = x^{-a}, \quad U(1, b, x) = x^{1-b} e^{x} \Gamma(b-1, x), \tag{A4.16}$$

where $\Gamma(\nu, x)$ is the upper incomplete gamma function. Using (A4.6) one obtains the formulas

Case 1: $BEG(\alpha, \delta, \beta) = BG(\upsilon = 0, \gamma = 1, \alpha, \delta, \beta)$

$$f(x) = \left(\tfrac{\beta}{\alpha+\beta}\right)^{\delta} \alpha e^{-\alpha x}, \quad x > 0, \quad f(x) = \Gamma(\delta)^{-1} \left(\tfrac{\beta}{\alpha+\beta}\right)^{\delta} \alpha e^{\alpha |x|} \cdot \Gamma(\delta, (\alpha+\beta)|x|), \quad x < 0. \tag{A4.17}$$

Case 2: $BGE(\gamma, \alpha, \beta) = BG(\upsilon = 0, \gamma, \alpha, \delta = 1, \beta)$

$$f(x) = \Gamma(\gamma)^{-1} \left(\tfrac{\alpha}{\alpha+\beta}\right)^{\gamma} \beta e^{\beta x} \cdot \Gamma(\gamma, (\alpha+\beta)x), \quad x > 0, \quad f(x) = \left(\tfrac{\alpha}{\alpha+\beta}\right)^{\gamma} \beta e^{-\beta|x|}, \quad x < 0. \tag{A4.18}$$

Bilateral exponential or skew Laplace $skL(\alpha, \beta) = BG(\upsilon = 0, \gamma = 1, \alpha, \delta = 1, \beta)$

$$f(x) = \left(\tfrac{\beta}{\alpha+\beta}\right) \alpha e^{-\alpha x}, \quad x > 0, \quad f(x) = \left(\tfrac{\alpha}{\alpha+\beta}\right) \beta e^{-\beta|x|}, \quad x < 0. \tag{A4.19}$$

Variance gamma $VG(\rho, \alpha, \beta) = BG(\upsilon = 0, \gamma = \rho, \alpha, \delta = \rho, \beta)$

In this situation the relevant Tricomi or Whittaker function reduces to a *Macdonald function* (modified Bessel function of the 2nd kind, hyperbolic Bessel function of the 3rd kind, Basset function, modified Hankel function) of the type (Atlas (2009), 48:4:3 and 48:13:6)

$$U(a, 2a, x) = \frac{x^{\frac{1}{2}-a}}{\sqrt{\pi}} e^{\frac{1}{2}x} K_{a-\frac{1}{2}}(\tfrac{1}{2}x), \quad W_{0,\mu}(x) = \sqrt{\frac{x}{\pi}} \cdot K_{\mu}(\tfrac{1}{2}x). \tag{A4.20}$$

Inserting these expressions into the Tricomi or Whittaker representation (A4.6) or (A4.14) one obtains the VG pdf

$$f(x) = \frac{(\alpha\beta)^{\rho}}{\sqrt{\pi}\Gamma(\rho)} \left(\frac{|x|}{\alpha+\beta}\right)^{\rho-\frac{1}{2}} \cdot \exp\left(-\tfrac{1}{2}(\alpha-\beta)x\right) \cdot K_{\rho-\frac{1}{2}}\left(\tfrac{1}{2}(\alpha+\beta)|x|\right), \quad x \neq 0. \tag{A4.21}$$

This closed-form expression has been first derived in Madan et al. (1998) for the parameterization

$$\left(\mu, \sigma^2, \nu\right) = \left((\alpha^{-1} - \beta^{-1})\rho, 2(\alpha\beta)^{-1}\rho, \rho^{-1}\right). \tag{A4.22}$$

In its original form the VG pdf takes the less symmetrical form

$$f(x) = \frac{2\exp(\sigma^{-2}\mu x)}{\nu^{\nu^{-1}} \sqrt{2\pi}\sigma \cdot \Gamma(\nu^{-1})} \cdot \left(\frac{x^2}{\mu^2 + 2\nu^{-1}\sigma^2}\right)^{\frac{1}{2}\nu^{-1}-\frac{1}{4}} \cdot K_{\nu^{-1}-\frac{1}{2}}\left(\sigma^{-2}\sqrt{(\mu^2 + 2\nu^{-1}\sigma^2)x^2}\right), \quad x \neq 0. \tag{A4.23}$$

In case the shape parameter $\gamma = \delta = \rho$ is an integer, further simplification to exponentials is obtained using the spherical Macdonald functions (Atlas (2009), Section 26:13).

Bilateral gamma with integer shape parameter

The case of integer shape parameter $\gamma = n$ or/and $\delta = m$ is best analysed in terms of the generalized gamma function representation of Theorem A4.1 and the following auxiliary result.

Lemma A4.1: (*Finite series expansion of the generalized gamma function*). For integer n one has

$$\Gamma(a,n,x) = \sum_{k=0}^{n-1} \frac{(a)_k (n-k)_k}{k} x^{-k} . \tag{A4.24}$$

Proof: Expanding the second term of the integral into a binomial series one obtains

$$\Gamma(a,n,x) = \int_0^\infty t^{a-1}(1+x^{-1}t)^{n-1} e^{-t} dt = \sum_{k=0}^{n-1} \binom{n-1}{k} \cdot x^{-k} \cdot \int_0^\infty t^{a+k-1} e^{-t} dt = \sum_{k=0}^{n-1} \frac{\Gamma(n)\Gamma(a+k)}{\Gamma(n-k)\Gamma(k+1)} x^{-k} .$$

Using that $\Gamma(k+1) = k\Gamma(k)$ and the Pochhammer symbol one obtains (A4.24). ◊

The following result relates to Proposition 3.1 in Küchler and Tappe (2008a), with the difference that our formulation is simpler and more symmetrical. It expresses the BG densities on each half of the real axis as finite linear combinations of gamma densities. The density of the standard gamma $\Gamma(\gamma,1)$ with scale parameter 1 is denoted by $g(x;\gamma) = \Gamma(\gamma)^{-1} x^{\gamma-1} e^{-x}$.

Theorem A4.2: (*Bilateral gamma density as finite linear combination of gamma densities*). For integer shape parameter $\gamma = n$ or/and $\delta = m$ the following gamma representations hold:

Case 1: $BG(\upsilon = 0, \gamma = n, \alpha, \delta, \beta), \quad x > 0$

$$f(x) = \left(\frac{\beta}{\alpha+\beta}\right)^\delta \cdot \sum_{k=0}^{n-1} \binom{\delta+k-1}{k} \left(\frac{\alpha}{\alpha+\beta}\right)^k \alpha \cdot g(\alpha x; n-k) \tag{A4.25}$$

Case 2: $BG(\upsilon = 0, \gamma, \alpha, \delta = m, \beta), \quad x < 0$

$$f(x) = \left(\frac{\alpha}{\alpha+\beta}\right)^\gamma \cdot \sum_{k=0}^{m-1} \binom{\gamma+k-1}{k} \left(\frac{\beta}{\alpha+\beta}\right)^k \beta \cdot g(\beta|x|; m-k) \tag{A4.26}$$

Proof: It suffices to insert (A4.24) into (A4.2) and rearrange terms taking into account the definition

$$\binom{a}{b} = \frac{\Gamma(a+1)}{\Gamma(b+1)\Gamma(a-b+1)}$$

of the generalized binomial coefficient. ◊

As a consequence the VG pdf with integer shape parameter is a closed-from expression. It suffices to restate (A4.25)-(A4.26) for $\gamma = \delta = n$ by changing the order of summation.

Variance gamma $VG(\rho = n, \alpha, \beta)$

$$\begin{aligned} f(x) &= \sum_{k=0}^{n-1} C_k \cdot \left(x^k e^{-\alpha x} \cdot 1\{x>0\} + |x|^k e^{-\beta|x|} \cdot 1\{x<0\}\right), \quad x \neq 0, \\ C_k &= \left(\frac{\alpha}{\alpha+\beta}\right)^n \left(\frac{\beta}{\alpha+\beta}\right)^n \binom{2(n-1)-k}{n-1} \frac{(\alpha+\beta)^{k+1}}{k!}. \end{aligned} \tag{A4.27}$$

Remarks A4.2: (A4.27) is also obtained by inserting into (A4.21) the following special case of the Macdonald function (Watson (1995), p.80, or Abramowitz and Stegun (1965), p.443):

$$K_{n+\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2x}} e^{-x} \cdot \sum_{k=0}^{n} \frac{(n+k)!}{k!(n-k)!} \frac{1}{(2x)^k}, \quad n = 0,1,2,... .$$

The VG with integer shape parameter is a member of the rich class of two-sided exponential-polynomial-trigonometric (EPT) functions, for which some interesting computational tools are available (see Sexton and Hanzon (2012), Section 3, and Hanzon et al. (2012), Section 7).

Example A4.2: NVG distribution

According to formula (2.1) with $\gamma = \delta = \rho$ the normal variance gamma $NVG(\upsilon, \tau, \rho, \alpha, \beta)$ is the convolution of the normal $N(\upsilon, \tau)$ and the variance gamma $VG(\rho, \alpha, \beta)$. Therefore, its distribution satisfies the infinite integral representation

$$F_{NVG}(x) = \int_{-\infty}^{\infty} \Phi(\tfrac{x-\upsilon-z}{\tau}) f_{VG}(z) dz, \quad \text{(A4.28)}$$

where $f_{VG}(x)$ is the pdf (A4.21) of the VG and $\Phi(x)$ is the standard normal distribution.

REFERENCES

1. *Aas, K. and H. Haff* (2006). The generalized hyperbolic skew Student's t-distribution. Journal of Financial Econometrics 4(2), 275-309.

2. *Abramowitz, M. and I. Stegun* (1965). Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. Dover Publications, New York.

3. *Anderson, T.W. and D.A. Darling* (1952). Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. Ann. Math. Statist. 23(2), 193–212

4. *Ardalan, A., Sadooghi-Alvandi, S.M. and A.R. Nematollahi* (2012). The two-piece normal-Laplace distribution. Comm. Statist. Theory Methods 41(20), 3759-3785.

5. *Arrow, K.* (1965). Aspects of the theory of risk bearing. Yrjö Jahnssonin Säätiö, Helsinki.

6. *Arrow, K.* (1971). Essays in the theory of risk bearing. Markham Economic Series, New York.

7. *Bellini, F. and L. Mercuri* (2012). Option pricing in a conditional bilateral gamma model. URL: http://ssrn.com/abstract=2046478

8. *Bera, A.K. and C.M. Jarque* (1987). A Test for Normality of Observations and Regression Residuals. Int. Statist. Rev. 55(2), 163-172.

9. *Bondesson, L.* (1992). Generalized gamma convolutions and related classes of distributions. Lecture Notes in Statistics 76. Springer-Verlag, Berlin.

10. *Briec, W., Kerstens, K and O. Jokung* (2007). Mean-variance-skewness portfolio performance gauging: a general shortage function and dual approach. Management Science 53(1), 135-149.

11. *Burnecki, K., Misiorek, A. and R. Weron* (2010). Loss Distributions. MPRA Paper No. 22163.

12. *Carr, P., Geman, H., Madan, D.B. and M. Yor* (2002). The fine structure of asset returns: an empirical investigation. J. Business 75(2), 305-332.

13. *Cesari, R. and D'adda* (2010). An ordinal utility of moments for decisions under uncertainty. Asian J. Math. Statist. 3(2), 52-68.

14. *Chan, T.* (1998). Some applications of Lévy processes to stochastic investment models for actuarial use. ASTIN Bulletin 28(1), 77-93.

15. *Cizek, P., Härdle, W. and R. Weron* (2005). Statistical Tools for Finance and Insurance. Springer.

16. *Conniffe, D.* (2007). The generalized extreme value distribution as utility function. Economic & Social Review 38(3), 275-288.

17. *Cramér, H.* (1928). On the composition of elementary errors: II. Skand. Akt. 11, 141-180.

18. *Cvitanic, J., Polimenis, V. and F. Zapatero* (2008). Optimal portfolio allocation with higher moments. Annals of Finance 4(1), 1-28.

19. *D'Agostino, R.B. and M.A. Stephens* (1986). Goodness-of-Fit Techniques. Marcel Dekker.

20. *Di Pierro, M. and Mosevich, J.* (2011). Effects of skewness and kurtosis on portfolio rankings. Quantitative Finance 11(10), 1449-1453.

21. *Domenig, Th. and P. Vanini* (2010). Stochastic volatility II – variance gamma model: analysis and implementation. URL: http://ssrn.com/abstract=1579923

22. *Fiorani, F.* (2004). Option pricing under the variance gamma process. Ph.D. Thesis. MPRA Paper no. 15395. URL: http://mpra.ub.uni-muenchen.de/15395/

23. *Fisher, R.A.* (1928). Moments and product moments of sampling distributions. Proc. London Math. Soc. 30, 199-238.

24. *Frecka, T. and W. Hopwood* (1983). The effects of outliers on the cross-sectional distributional properties of financial ratios. The Accounting Review 58(1), 115-128.

25. *Fu, M.C., Jarrow, R.A., Yen, J-Y. and R.J. Elliot* (2006). Advances in Mathematical Finance (Festschrift D. Madan 60th birthday). Appl. Numerical Harmonic Analysis, special issue.

26. *George, F.* (2007). Johnson's system of distributions and microarray data analysis. Ph.D. Thesis. Graduate School Theses and Dissertations. Paper 2186. URL: http://scholarcommons.usf.edu/etd/2186/

27. *Geman, H.* (2002). Pure jump Lévy processes for asset price modelling. J. of Banking and Finance 26, 1297-1316.

28. *Ghysels, E. and F. Wang* (2011). Some useful densities for risk management and their properties. Preprint, forthcoming in Econometric Reviews. URL: http://www.unc.edu/~eghysels/working_papers.html

29. *Gradshteyn, I.S. and I.M. Ryzhik* (2000). Table of integrals, series and products. Academic Press, San Diego.

30. *Guiard, V.* (1980). Robustheit I. Probleme der angewandten Statistik, Heft 4, FZ für Tierproduktion Dummerstorf-Rostock.

31. *Hansen, B.E.* (1994). Autoregressive conditional density estimation. International Economic Review 35(3), 705-730.

32. *Hanzon, B., Holland, F. and C. Sexton* (2012). Financial modelling with 2-EPT Lévy processes. URL: http://www.2-ept.com/

33. *Hlawitschka, W.* (1994). The empirical nature of Taylor-series approximations to expected utility. American Economic Review 84, 713-719.

34. *Hürlimann, W.* (2008a). The Normal Laplace approximation to compound distributions. Journal of Pure and Applied Mathematics 44(3), 415-424.

35. *Hürlimann, W.* (2008b). Extremal Moment Methods and Stochastic Orders – Application in Actuarial Science. Boletín de la Asociación Matemática Venezolana (BAMV), Volume XV, num.1 & num. 2.

36. *Hürlimann, W.* (2009). Robust variants of Cornish-Fischer approximation and Chebyshev-Markov bounds: application to value-at-risk. Adv. Appl. in Math. Sci. 1(2), 239-260.

37. *Hürlimann, W.* (2012). On non-Gaussian AR(1) inflation modelling. J. Statistical and Econometric Methods 1(1), 93-109.

38. *Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U. and T.P. Speed* (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4(2), 249-264.

39. *Johnson, N.L.* (1949). System of frequency curves generated by methods of translation. Biometrika 36, 149-176.

40. *Johnson, N.L., Kotz, S. and N. Balakrishnan* (1994). Continuous Univariate Distributions, vol. 1 (2nd ed.). J. Wiley, New York.

41. *Jondeau, E. and M. Rockinger* (2003). Conditional volatility, skewness and kurtosis: existence, persistence and comovements. J. Econom. Dynam. Control 27, 1699-1737.

42. *Jondeau, E. and M. Rockinger* (2006). Optimal portfolio allocation under higher moments. European Financial Management 12(1), 29-55.

43. *Jurczencko, E.F., Maillet, B.B. and P.M. Merlin* (2008). Efficient frontier for robust higher-order moment portfolio selection. URL: ftp://mse.univ-paris1.fr/pub/mse/CES2008/Bla08062.pdf

44. *Kaishev, V.K.* (2010). Stochastic process induces by Dirichlet (B-) splines: modelling multivariate asset price dynamics. Actuarial Research paper no. 195, Cass Business School, London.

45. *Kallbergaard, J.G. and W.T. Ziemba* (1979). On the robustness of the Arrow-Pratt risk aversion measure. Economics Letters 2, 21-26.

46. *Kleniati, P.M. and B. Rustem* (2009). Portfolio decisions with higher moments. COMISEF Working Paper Series, WPS-021. URL: www.comisef.eu

47. *Kotz, S., Kozubowski, T.J. and K. Podgorski* (2001). The Laplace distribution and generalizations: a revisit with applications. Birkhäuser, Boston.

48. *Koponen, I.* (1995). Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process. Physical Review E 52, 1197-1199.

49. *Küchler, U. and S. Tappe* (2008a). On the shapes of bilateral gamma densities. Statist. Probab. Lett. 78(15), 2478-2484.

50. *Küchler, U. and S. Tappe* (2008b). Bilateral gamma distributions and processes in financial mathematics. Stochastic Process. Appl. 118(2), 261-283.

51. *Küchler, U. and S. Tappe* (2009). Option pricing in bilateral gamma stock models. Statistics and Decisions 27, 281-307.

52. *Kundu, D.* (2005). Discriminating between the normal and Laplace distributions. In: N. Balakrishnan, H.N. Nagaraja and N. Kannan (eds.). Advances in Ranking and Selection, Multiple Comparisons, and Reliability, 65-78. Birkhäuser, Boston.

53. *Lévy, H. and H. Markowitz* (1979). Approximating expected utility by a function of mean and variance. Annals of Econom. Review 69, 308-317.

54. *Lishamol, T. and K.K. Jose* (2009). Generalized normal-Laplace AR processes. Statist. Probab. Lett. 79(14), 1615-1620.

55. *Madan, D.* (2001). Purely discontinuous asset pricing processes. In: Jouini, E., Cvitanic, J. and M. Musiela (Eds.). Option Pricing, Interest Rates and Risk Management, 105-153. Cambridge University Press, Cambridge.

56. *Madan, D. and E. Seneta* (1990). The variance gamma model for share market returns. Journal of Business 63, 511-524.

57. *Madan, D., Carr, P and E. Chan*g (1998). The variance gamma process and option pricing. European Finance Review 2, 79-105.

58. *Madan, D. and F. Milne* (1991). Option pricing with VG martingale components. Mathematical Finance 1(4), 3955.

59. *Markowitz, H.M.* (1952). Portfolio selection. Journal of Finance 7(1), 77-91.

60. *Mathai, A.M.* (1993a). On non-central generalized Laplacianness of quadratic forms in normal

61. *Mathai, A.M.* (1993b). The residual effect of a growth-decay mechanism and the distributions of covariance structures. Canad. J. Statist.21 (3), 227-283.

62. *Merton, R.C.* (1971). Optimum consumption and portfolio rules in a continuous time model. Journal of Economic Theory 3, 373-413.

63. *Mises, R. von* (1931). Wahrscheinlichkeitsrechnung und Ihre Anwendung in der Statistik und Theoretischen Physik. Deuticke, Leipzig. variables. J. Multivariate Anal.45 (2), 239-246.

64. *Nash, J.F.* (1950). The bargaining problem. Econometrica 18, 155.

65. *Neumann, J. von and O. Morgenstern* (1947). Theory of Games and Economic Behavior (2nd ed.). Princeton University Press, Princeton, New Jersey.

66. *Oldham, K., Myland, J. and J. Spanier* (2009). An Atlas of Functions – with Equator, the Atlas Function Calculator (2nd ed.). Springer, New York.

67. *Pearson, K.* (1916). Mathematical contributions to the theory of evolution XIX; second suppl. to a memoir on skew variation. Phil. Trans. Royal Soc. London, Ser. A 216, 432.

68. *Prakash, A.J., Chang, C.-H., and T.E. Pactwa* (2003). Selecting a portfolio with skewness: recent evidence from US, European, and Latin American equity markets. J. Banking Finance 27, 1375-1390.

69. *Plancade, S., Rozenholc, Y. and E. Lund* (2011). Improving background correction for Illumina BeadArrays: the normal-gamma model. URL: http://arxiv.org/pdf/1112.4180.pdf

70. *Pratt, J.W.* (1964). Risk aversion in the small and in the large. Econometrica 32, 122-136.

71. *Pulley, L.B.* (1981). A general mean-variance approximation to expected utility for short holding periods. Journal of Financial and Quantitative Analysis16, 361-373.

72. *Reed, W.J.* (2006). The normal Laplace distribution and its relatives. In: Balakrishnan, N., Castillo, E. and J.M. Sarabia (Eds.). Advances in Distribution Theory, Order Statistics and Inference. Statistics for Industry and Technology. Birkhäuser.

73. *Reed, W.J.* (2007). Brownian-Laplace motion and its use in financial modelling. Comm. Statist. Theory Methods 36, 473-484.

74. *Reed, W.J.* (2011). A flexible parametric survival model which allows a bathtub shaped hazard rate function. J. Appl. Statistics 38(8), 1665-1680.

75. *Reed, W.J. and M. Jorgensen* (2004). The double Pareto-lognormal distribution – a new parametric model for size distributions. Comm. Statist. Theory Methods 33(8), 1733-53.

76. *Reed, W.J. and A. Pewsey* (2009). Two nested families of skew-symmetric circular distributions.TEST 18(3), 516-528.

77. *Samuelson, P.* (1970). The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. Rev. Econom. Studies 37(1-2), 537-542.

78. *Sato, K.* (1999). Lévy processes and infinitely divisible distributions. Cambridge Studies in Advanced Mathematics, Cambridge.

79. *Sexton, C. and B. Hanzon* (2012). State space calculations for two-sided EPT densities with financial modelling applications. URL: http://www.2-ept.com/

80. *Skoulakis, G.* (2012). On the quality of Taylor approximations to expected utility. Applied Financial Economics 22(11), 863-876.

81. *Stein, H.A., Hogan, P. and P. Carr* (2007). Time for a change: the variance gamma model and option pricing. Working Paper, Bloomberg Financial Markets, New York. URL: http://ssrn.com/abstract=956625

82. *Theodossiu, P.* (1998). Financial data and the skewed generalized T distribution. Management Science 44(12), 1650-1661.

83. *Thorin, O.* (1978). An extension of the notion of a generalized Γ-convolution. Scand. Act. J., 141-149.

84. *Tricomi, F.G.* (1947). Sulle funzioni ipergeometriche confluenti (in Italian). Annali di Matematica Pura ed Applicata. Serie Quarta 26: 141–175.

85. *Watson, G.N.* (1995). A Treatise on the Theory of Bessel Functions. Cambridge Mathematical Library, Cambridge.

86. *Wichitaksorn, N., Choy, S.T.B. and R. Gerlach* (2012). A generalized class of skew distributions and parametric quantile regression models. URL: http://ssrn.com/abstract=2122378.

87. *Wilkins, J.E.* (1944). A note on skewness and kurtosis. Annals Math. Statist. 15, 333-335.

88. *Xie, Y., Wang, X. and M. Story* (2009). Statistical methods of background correction for illumina beadarray data. Bioinformatics 25(6), 751-757.

89. *Yu, K. and J. Zhang* (2005). A three-parameter asymmetric Laplace distribution and its extension. Comm. Statist. Theory Methods 34(9-10), 1867-1879. Erratum (2006). Comm. Statist. Theory Methods 35(7-9), 1359.

Source of support: Nil, Conflict of interest: None Declared