# The Generalized Hyperbolic Model:
# Estimation, Financial Derivatives, and Risk Measures

Dissertation zur Erlangung des Doktorgrades
der Mathematischen Fakultät
der Albert-Ludwigs-Universität Freiburg i. Br.

vorgelegt von

## Karsten Prause

Oktober 1999

*To Ulrike*

# Preface

The aim of this dissertation is to describe more realistic models for financial assets based on generalized hyperbolic (GH) distributions and their subclasses.

Generalized hyperbolic distributions were introduced by Barndorff-Nielsen (1977), and stochastic processes based on these distributions were first applied by Eberlein and Keller (1995) in Finance. Being a normal variance-mean mixture, GH distributions possess semi-heavy tails and allow for a natural definiton of volatility models by replacing the mixing generalized inverse Gaussian (GIG) distribution by appropriate volatility processes.

In the first Chapter we introduce univariate GH distributions, construct an estimation algorithm and examine statistically the fit of generalized hyperbolic distributions to log-return distributions of financial assets. We extend the hyperbolic model for the pricing of derivatives to generalized hyperbolic Lévy motions and discuss the calculation of prices by fast Fourier methods and saddle-point approximations.

Chapter 2 contains on the one hand a general recipe for the evaluation of option pricing models; on the other hand the derivative pricing based on GH Lévy motions is studied from various points of view: The accordance with observed asset price processes is investigated statistically, and by simulation studies the sensitivity to relevant variables; finally, theoretical prices are compared with quoted option prices. Furthermore, we examine two approaches to martingale modelling and discuss alternative ways to test option pricing models.

Barndorff-Nielsen (1998) proposed a refinement of the GH Lévy model by replacing the mixing GIG distribution by a volatility process of the Ornstein-Uhlenbeck type. We investigate this model in Chapter 3 with a view towards derivative pricing. After a review of this model we derive the minimal martingale measure to compute option prices and investigate the behaviour of this model thoroughly from a numerical and econometric point of view.

We start in Chapter 4 with a description of some "stylized features" observed in multivariate return distributions of financial assets. Then we introduce multivariate GH distributions and discuss the efficiency of estimation procedures. Finally, the multivariate Esscher approach to option pricing and, in particular, basket options are examined.

In the final chapter we define more realistic risk measures based on generalized hyperbolic distributions and evaluate them following the procedure required by the Basel Committee on Banking Supervision. The proposed risk-measurement methods apply to linear and nonlinear portfolios; moreover they are computationally not demanding.

Forecasting values of financial assets is not a major objective of this dissertation. However, the second and the final chapter presents some results on volatility forecasts for option prices and on tests of risk measures based on distributional forecasts.

The usual procedure will always be to start with the definition of the new model and the calibration of the model. Then the properties of the model are examined theoretically and by simulation studies. Finally, we will put the model into practice and compare the performance

with benchmark models. This procedure yields a complete picture and highlightens features relevant for practical applications. The final two chapters are self-contained. Therefore, a reader only interested in multivariate modelling could skip the preceeding three chapters.

# Contents

# Chapter 1

# Generalized Hyperbolic Lévy Motions

The modelling of financial assets as stochastic processes is determined by distributional assumptions on the increments and the dependence structure. It is well known that the returns of most financial assets have semi-heavy tails, i.e. the actual kurtosis is higher than the kurtosis of the normal distribution (Mandelbrot 1963). Generalized hyperbolic distributions (GH) possess these semi-heavy tails. On the contrary, the use of models with nonexisting moments, e.g. stable distributions, does not admit the pricing of derivative contracts by martingale methods where we aim at. Note, that in contrast to stable distributions the density of generalized hyperbolic distributions is known explicitly.

Ole E. Barndorff-Nielsen (1977) introduced the generalized hyperbolic distributions and at first applied them to model grain size distributions of wind blown sands. An important aspect is, that GH distributions embrace many special cases, respectively limiting distributions, of hyperbolic, normal inverse Gaussian (NIG), Student-t, variance-gamma and normal distributions. All of them have been used to model financial returns. The mathematical properties of these univariate distributions are well-known (Barndorff-Nielsen and Blæsild 1981; Blæsild 1999). Recently, generalized hyperbolic distributions resp. their subclasses have been proposed as a model for the distribution of increments of financial price processes (Eberlein and Keller 1995; Rydberg 1997a; Barndorff-Nielsen 1998; Eberlein, Keller, and Prause 1998) and as limit distributions of diffusions (Bibby and Sørensen 1997). Jaschke (1997) remarked that one could also find a stationary generalized hyperbolic distribution as a weak limit of a process with stochastic volatility, which is modelled by a GARCH(1,1) time series.

In this chapter we present new estimation results for German stock, US stock index and high-frequency data, and investigate the goodness of fit. In particular, we look at the tails of the distributions. In the final sections we extend the option pricing approach based on the Esscher transform, proposed by Eberlein and Keller (1995), Keller (1997) to generalized hyperbolic distributions and discuss the numerical problems.

## 1.1 Generalized Hyperbolic Distributions

We start with an exposition of the univariate generalized hyperbolic distributions and the subclasses which are relevant for applications.

### Definition and Parametrizations

**Definition 1.1.** *The one-dimensional* generalized hyperbolic *(GH) distribution is defined by the following Lebesgue density*

$$\mathrm{gh}(x; \lambda, \alpha, \beta, \delta, \mu) = a(\lambda, \alpha, \beta, \delta) \left(\delta^2 + (x-\mu)^2\right)^{(\lambda-\frac{1}{2})/2}$$
$$\times \mathrm{K}_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) \exp(\beta(x-\mu))$$
$$a(\lambda, \alpha, \beta, \delta) = \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi}\,\alpha^{\lambda-1/2}\,\delta^\lambda\,\mathrm{K}_\lambda\left(\delta\sqrt{\alpha^2 - \beta^2}\right)}$$

*where* $\mathrm{K}_\lambda$ *is a modified Bessel function and* $x \in \mathbb{R}$. *The domain of variation of the parameters is* $\mu \in \mathbb{R}$ *and*

$$\delta \geq 0,\ |\beta| < \alpha \qquad \textit{if } \lambda > 0$$
$$\delta > 0,\ |\beta| < \alpha \qquad \textit{if } \lambda = 0$$
$$\delta > 0,\ |\beta| \leq \alpha \qquad \textit{if } \lambda < 0.$$

Different scale- and location-invariant parametrizations of the generalized hyperbolic distribution have been proposed in the literature.

| | | |
|---|---|---|
| 2nd parametrization | $\zeta = \delta\sqrt{\alpha^2 - \beta^2},\ \varrho = \beta/\alpha$ | (1.2) |
| 3rd parametrization | $\xi = (1 + \zeta)^{-1/2},\ \chi = \xi\varrho$ | (1.3) |
| 4th parametrization | $\bar{\alpha} = \alpha\delta,\ \bar{\beta} = \beta\delta$ | (1.4) |

Note, that for symmetric distributions $\beta = \bar{\beta} = \varrho = \chi = 0$ holds. Rewriting the density of GH distributions in the 4th parametrization, we obtain a representation in which the role of the parameters $\delta$ and $\mu$ describing scale and location is more obvious. The density above and those for the hyperbolic and NIG distributions in the 4th parametrization are presented in Appendix C. Scale- and location-invariant parameters are given in the following lemma.

**Lemma 1.5.** *The terms* $\lambda$, $\bar{\alpha}$ *and* $\bar{\beta}$ *are scale- and location-invariant parameters of the univariate generalized hyperbolic distribution. The very same holds for the other parametrizations* $(\zeta, \varrho)$ *and* $(\xi, \chi)$ *defined in (1.2) and (1.3).*

*Proof.* Blæsild (1981, Theorem I) has proved that a linear transformation $Y = aX + b$ of $X \sim \mathrm{GH}_1$ is again GH-distributed with parameters $\lambda^+ = \lambda$, $\alpha^+ = \alpha/|a|$, $\beta^+ = \beta/|a|$, $\delta^+ = \delta|a|$ and $\mu^+ = a\mu + b$. Obviously $\alpha^+\delta^+ = \alpha\delta$ and $\beta^+\delta^+ = \beta\delta$ holds. $\square$

The parameters $\mu$ and $\delta$ describe the location and the scale, whereas $\beta$ describes the skewness. Increasing $\xi$ or decreasing $\zeta$ or $\bar{\alpha}$ reflect an increase in the kurtosis. Note, that

generalized hyperbolic distributions can also be represented as a normal variance-mean mixture

$$\mathrm{gh}(x; \lambda, \alpha, \beta, \delta, \mu) = \int_0^\infty N(x; \mu + \beta w, w) \, \mathrm{gig}(w; \lambda, \delta^2, \alpha^2 - \beta^2) dw, \qquad (1.6)$$

where $N$ is the normal density function with respect to mean and variance. Furthermore, $\mathrm{gig}(x; \lambda, \chi, \psi)$ denotes the density function of generalized inverse Gaussian distributions.

**Definition 1.7.** *The* generalized inverse Gaussian *(GIG) distribution is given by the Lebesgue density*

$$\mathrm{gig}(x; \lambda, \chi, \psi) = \frac{(\psi/\chi)^{\lambda/2}}{2 \, \mathrm{K}_\lambda(\sqrt{\psi\chi})} \, x^{\lambda-1} \exp\Big( - \frac{1}{2}\big(\chi x^{-1} + \psi x\big) \Big), \quad x > 0,$$

*where $\lambda \in \mathrm{I\!R}$ and $\chi, \psi \in \mathrm{I\!R}_+$.*

The standard reference for the GIG distribution is Jørgensen (1982). In Figure 3.2 we show the GIG distributions implied as *mixing distributions* in the univariate and multivariate GH distributions.

## Limiting Distributions and Subclasses

**Remark 1.8.** *The normal distribution is obtained as a limiting case of the generalized hyperbolic distribution for $\delta \to \infty$ and $\delta/\alpha \to \sigma^2$ (see also for other limiting distributions Barndorff-Nielsen 1978).*

Using the properties of Bessel functions $\mathrm{K}_\lambda$, it is possible to simplify the density gh whenever $\lambda \in {}^1\!/_2 \, \mathbb{Z}$. For $\lambda = n + {}^1\!/_2$, $n = 0, 1, 2, \ldots$ the Bessel function $\mathrm{K}_\lambda$ may be expressed as

$$\mathrm{K}_{n+{}^1\!/_2}(x) = \sqrt{\frac{\pi}{2}} \, x^{-1/2} \, e^{-x} \left( 1 + \sum_{i=1}^n \frac{(n+i)!}{(n-i)! \, i!} (2x)^{-i} \right). \qquad (1.9)$$

Since $\mathrm{K}_\lambda(x) = \mathrm{K}_{-\lambda}(x)$, we obtain $\mathrm{K}_{-1/2}(x) = \mathrm{K}_{1/2}(x) = \sqrt{\pi/2} \, x^{-1/2} \, e^{-x}$. In particular, this equation allows to deduce simpler expressions in the cases $\lambda = -1/2$ and $\lambda = 1$.

**Definition 1.10.** *For $\lambda = 1$ we obtain* hyperbolic distributions *(HYP)*

$$\mathrm{hyp}(x; \alpha, \beta, \delta, \mu) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\delta\alpha \, \mathrm{K}_1\big(\delta\sqrt{\alpha^2 - \beta^2}\big)} \exp\big( - \alpha\sqrt{\delta^2 + (x - \mu)^2} + \beta(x - \mu) \big),$$

*where $x, \mu \in \mathrm{I\!R}$, $0 \leq \delta$ and $|\beta| < \alpha$.*

**Definition 1.11.** *For $\lambda = -1/2$ we get the* normal inverse Gaussian *(NIG) distribution with Lebesgue density*

$$\mathrm{nig}(x; \alpha, \beta, \delta, \mu) = \frac{\alpha\delta}{\pi} \exp\Big(\delta\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)\Big) \frac{\mathrm{K}_1\big(\alpha\sqrt{\delta^2 + (x - \mu)^2}\big)}{\sqrt{\delta^2 + (x - \mu)^2}},$$

*where $x, \mu \in \mathrm{I\!R}$, $0 \leq \delta$ and $0 \leq |\beta| \leq \alpha$.*

| GH<br>Generalized hyperbolic | GIG<br>Generalized inverse Gaussian | Mixing GIG distribution | | |
| --- | --- | --- | --- | --- |
| | | $\lambda$ | $\chi$ | $\psi$ |
| General case (Blæsild 1999) | General case: (Barndorff-Nielsen and Shephard 1998, 4.6.1) | $\in \mathbb{R}$ | $\delta^2$ | $\alpha^2 - \beta'\Delta\beta$ |
| Hyperbolic (Eberlein and Keller 1995) | Positive Hyperbolic (Barndorff-Nielsen and Shephard 1998, 4.6.3) | $(d+1)/2$ | $\delta^2$ | $\alpha^2 - \beta'\Delta\beta$ |
| Hyperboloid | | $(d-1)/2$ | $\delta^2$ | $\alpha^2 - \beta'\Delta\beta$ |
| Normal inverse Gaussian (NIG) | Inverse Gaussian (IG) | $-1/2$ | $\delta^2$ | $\alpha^2 - \beta'\Delta\beta$ |
| Normal Reciprocal Inverse Gaussian (NRIG) | Reciprocal Inverse Gaussian (RIG) | $1/2$ | $\delta^2$ | $\alpha^2 - \beta'\Delta\beta$ |
| Normal $\mathrm{N}^d(\mu + w_0\beta\Delta, w_0\Delta)$ in the case $\alpha, \delta \to \infty$ and $\delta/\alpha \to w_0$ (Barndorff-Nielsen 1978; Blæsild 1999) | | $\in \mathbb{R}$ | $\to \infty$ | $\to \infty$ |
| Variance gamma $(\sigma, \nu, \vartheta)$ is GH with $\lambda = \sigma^2/\nu$, $\alpha = \sqrt{(2/\nu) + (\vartheta^2/\sigma^4)}$, $\beta = \vartheta/\sigma^2$, $\delta = 0$ and $\mu = 0$ (Madan, Carr, and Chang 1998) | Gamma distribution $\Gamma(\lambda, \psi/2)$ (Wirth 1998; Barndorff-Nielsen and Shephard 1998, 4.6.5 and Blæsild 1999; see also Section 3.2) | $> 0$ | $0$ | $> 0$ |
| Student $t_d(\nu, \mu)$, $\nu \in \mathbb{R}_+$ is GH with with $\lambda = -\nu/2$, $\alpha = \beta = 0$, $\delta = \sqrt{\nu}$ (Blæsild 1999) | Inverse Gamma (inverse chi-squared) with $\nu \in \mathbb{R}_+$ degrees of freedom: $\Gamma^{-1}(\nu, 1/2)$, (Barndorff-Nielsen and Shephard 1998, 4.6.4) | $-\nu/2 < 0$ | $\nu > 0$ | $0$ |
| $\mathrm{Cauchy}_d(\mu)$ is GH with $\lambda = -1/2$, $\alpha = \beta = 0$, $\delta = 1$ (Blæsild 1999) | | $-1/2$ | $1$ | $0$ |
| $\mathrm{GIG}(\lambda, \tau, \psi)$ in the case $\alpha\delta^2 \to \tau$, $\alpha - \beta = \psi/2$ and $\mu = 0$ (Blæsild 1999) | | $\in \mathbb{R}$ | $\to 0$ | const. |
| Skewed Laplace, $\lambda = (d+1)/2$, $\delta = 0$ and $\Delta = I_d$ (Blæsild 1999) | | $(d+1)/2$ | $0$ | $\alpha^2 - \beta'\Delta\beta$ |

Table 1.1: *Generalized hyperbolic as variance-mean mixtures of generalized inverse Gaussian distributions. Barndorff-Nielsen and Shephard (1998) use the following notation $\lambda = \bar{\lambda}$, $\chi = \delta^2$ and $\psi = \gamma^2$ for the GIG distribution; $d$ denotes the dimensions.*

A detailed description of the NIG distribution is given in Rydberg (1997a) and Barndorff-Nielsen (1997, 1998). See Barndorff-Nielsen and Prause (1999) for an application of NIG distributions to high-frequency financial data.

**Remark 1.12.** *The Student-t distribution results from a mixture of normal and inverse gamma distributions, which are obtained in the case $\psi = 0$ and $\lambda < 0$ from the GIG distributions. Therefore, we have a Student-t distribution as a limit of GH distributions for $\lambda < 0$ and $\alpha = \beta = \mu = 0$ (Barndorff-Nielsen 1978).*

Table 1.1 contains an overview over subclasses and limiting distributions of GH and the corresponding mixing GIG distributions. All mentioned distributions are characterized as variance-mean mixtures with special cases of GIG distributions.

## Properties

First we calculate the moment generating function which we also need in the Esscher approach to option pricing. Moment generating functions are also crucial for the the saddle-point approximation to compute option prices and quantiles of convolutions.

**Lemma 1.13.** *The moment generating function of the generalized hyperbolic distribution is given by*

$$M(u) = e^{u\mu} \left( \frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + u)^2} \right)^{\lambda/2} \frac{\mathrm{K}_\lambda\big(\delta\sqrt{\alpha^2 - (\beta + u)^2}\big)}{\mathrm{K}_\lambda\big(\delta\sqrt{\alpha^2 - \beta^2}\big)}, \quad |\beta + u| < \alpha.$$

The moment generating functions for hyperbolic and NIG distribution are obtained as special cases for $\lambda = 1$ and $\lambda = -1/2$ respectively.

*Proof.* We assume that $\mu = 0$. The norming factor $a$ is given in Definition 1.1. Then we obtain for $|\beta + u| < \alpha$

$$M(u) = \int e^{ux} \mathrm{gh}(x; \lambda, \alpha, \beta, \delta, 0) dx$$

$$= a(\lambda, \alpha, \beta, \delta) \int e^{ux} (\delta^2 + x^2)^{\frac{1}{2}(\lambda - \frac{1}{2})} \mathrm{K}_{\lambda - \frac{1}{2}}(\alpha\sqrt{\delta^2 + x^2}) e^{\beta x} dx$$

$$= \frac{a(\lambda, \alpha, \beta, \delta)}{a(\lambda, \alpha, \beta + u, \delta)}$$

$$= \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi}\delta^\lambda \alpha^{\lambda - \frac{1}{2}} \mathrm{K}_\lambda(\delta\sqrt{\alpha^2 - \beta^2})} \frac{\sqrt{2\pi}\delta^\lambda \alpha^{\lambda - \frac{1}{2}} \mathrm{K}_\lambda(\delta\sqrt{\alpha^2 - (\beta + u)^2})}{(\alpha^2 - (\beta + u)^2)^{\lambda/2}}$$

$$= \left( \frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + u)^2} \right)^{\lambda/2} \frac{\mathrm{K}_\lambda(\delta\sqrt{\alpha^2 - (\beta + u)^2})}{\mathrm{K}_\lambda(\delta\sqrt{\alpha^2 - \beta^2})};$$

the result follows since $\mu$ is the location parameter. $\qquad\Box$

Now we can calulate the mean and variance of the generalized hyperbolic distribution by differentiating the moment generating function.

**Proposition 1.14 (Mean and Variance).** *The generalized hyperbolic distribution has the following mean and variance*

$$\mathrm{E}\,X \;\; = \;\; \mu + \frac{\beta\delta}{\sqrt{\alpha^2 - \beta^2}}\frac{\mathrm{K}_{\lambda+1}(\zeta)}{\mathrm{K}_\lambda(\zeta)} \tag{1.15}$$

$$\mathrm{Var}\,X \;\; = \;\; \delta^2\left(\frac{\mathrm{K}_{\lambda+1}(\zeta)}{\zeta\,\mathrm{K}_\lambda(\zeta)} + \frac{\beta^2}{\alpha^2 - \beta^2}\left[\frac{\mathrm{K}_{\lambda+2}(\zeta)}{\mathrm{K}_\lambda(\zeta)} - \left(\frac{\mathrm{K}_{\lambda+1}(\zeta)}{\mathrm{K}_\lambda(\zeta)}\right)^2\right]\right) \tag{1.16}$$

*where $\zeta = \delta\sqrt{\alpha^2 - \beta^2}$. In (1.16) the term in the round brackets is scale- and location-invariant.*

*Proof.* With Gut (1995, Theorem III.3.3) all moments of $X$ exist. We apply the equation $\mathrm{K}'_\lambda(x) = -\,\mathrm{K}_{\lambda+1}(x) + \frac{\lambda}{x}\,\mathrm{K}_\lambda(x)$ for the derivative of the Bessel function and assume $\mu = 0$ without loss of generality.

$$
\begin{aligned}
M'(u) \;\; &= \;\; \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\mathrm{K}_\lambda(\delta\sqrt{\alpha^2 - \beta^2})}\left(\mathrm{K}_\lambda\big(\delta\sqrt{\alpha^2 - (\beta+u)^2}\big)\big(\alpha^2 - (\beta+u)^2\big)^{-\lambda/2}\right)' \\[2mm]
&= \;\; \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\mathrm{K}_\lambda(\delta\sqrt{\alpha^2 - \beta^2})}\frac{\mathrm{K}_{\lambda+1}\big(\delta\sqrt{\alpha^2 - (\beta+u)^2}\big)}{(\alpha^2 - (\beta+u)^2)^{(\lambda+1)/2}}
\end{aligned}
\tag{1.17}
$$

If we insert 0 for $u$ we obtain

$$M'(0) = \frac{\beta\delta\,\mathrm{K}_{\lambda+1}(\delta\sqrt{\alpha^2 - \beta^2}))}{\sqrt{\alpha^2 - \beta^2}\,\mathrm{K}_\lambda(\delta\sqrt{\alpha^2 - \beta^2}))}$$

Applying the definition of $\zeta$ gives the result for the mean; the similar computation of the variance is more complicated but straightforward. The scale- and location-invariance of the term in the round brackets of equation (1.16) follows with Lemma 1.5. $\qquad\square$

Clearly, both formulae are less complicated in the symmetric case, e.g. when we assume $\beta = 0$ the mean is simply $\mu$.

**Lemma 1.18.** *The characteristic function of the generalized hyperbolic distribution is given by*

$$\varphi(u) = e^{i\mu u}\left(\frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta+iu)^2}\right)^{\lambda/2}\frac{\mathrm{K}_\lambda\big(\delta\sqrt{\alpha^2 - (\beta+iu)^2}\big)}{\mathrm{K}_\lambda\big(\delta\sqrt{\alpha^2 - \beta^2}\big)}.$$

*Proof.* The radius of convergence of the moment generating function $M$ around zero is $\alpha - \beta$. With Gut (1995, Theorem III.3.3) the moment generating function $M$ is a real analytic, i.e. it can be expanded in a power series around zero. Consequently $M$ is a holomorphic function for complex $z$ with $|z| < \alpha - \beta$. Therefore, we obtain for the characteristic function $\varphi(u) = M(iu)$. $\qquad\square$

## Shape Triangle

In the case of hyperbolic distributions the scale- and location-invariant parameters $(\chi, \xi)$ are used to determine the shape of the distribution, because $(\chi, \xi)$ describe asymptotically

skewness and kurtosis (Barndorff-Nielsen, Blæsild, Jensen, and Sørensen 1985): for not too large values of $\varrho$ we have the approximations for the skewness $\gamma_1 \approx 3\chi$ and kurtosis $\gamma_2 \approx 3\xi^2$. The domain of variation of GH distributions given in the 3rd parametrization is $0 \leq |\chi| < \xi < 1$. Hence, they may be plotted in a triangle. This leads to a useful visualization of the shape in the *shape triangle* $(\xi, \chi)$.

It is possible to extent the concept of the shape triangle to generalized hyperbolic distributions with fixed $\lambda$, but the approximations of skewness and kurtosis are not valid in general. Hence we also compare the values of skewness and kurtosis directly.

See Eberlein and Keller (1995), Barndorff-Nielsen and Prause (1999) for an application of the hyperbolic and NIG shape triangle to analyze returns of financial assets on different time scales. In the Figures 2.24 and 5.12 we use shape trianglees to compare risk-neutral distributions and return distributions of nonlinear portfolios respectively.

We often restrict our empirical investigations to these two subclasses with fixed $\lambda$ because hyperbolic distributions are the fastest to estimate (see next section) and NIG distributions are closed under convolution.

### Tail Behaviour

The generalized hyperbolic distributions have semi-heavy tails, in particular

$$\text{gh}(x; \lambda, \alpha, \beta, \delta) \sim |x|^{\lambda-1} \exp\big((\mp\alpha + \beta)x\big) \text{ as } x \to \pm\infty \tag{1.19}$$

up to a multiplicative constant (Barndorff-Nielsen and Blæsild 1981, equation 15).

## 1.2  Maximum-Likelihood Estimation

In this section we begin with a description of the ML estimation algorithm: We assume the independence of the observations $x_i$, $i = 1, \ldots, n$ and maximize the log-likelihood function

$$L = \log a(\lambda, \alpha, \beta, \delta) + \left(\frac{\lambda}{2} - \frac{1}{4}\right) \sum_{i=1}^{n} \log\left(\delta^2 + (x - \mu)^2\right) \tag{1.20}$$

$$+ \sum_{i=1}^{n} \left[ \log \mathrm{K}_{\lambda-\frac{1}{2}}\big(\alpha\sqrt{\delta^2 + (x_i - \mu)^2}\big) + \beta(x_i - \mu)\right],$$

where $a$ is given in Definition 1.1. For hyperbolic ($\lambda = 1$), NIG ($\lambda = -1/2$), hyperbola ($\lambda = 0$), and hyperboloid ($\lambda = 1/2$) distributions the algorithm uses the simpler expressions of the log-likelihood function. We have chosen a numerical estimation procedure mainly based on an optimization for each coordinate. For the optimization step in each direction we have implemented a refined bracketing method (Jarrat 1970; Thisted 1988) which makes no use of derivatives. This gives us the possibility to replace the log-likelihood function easily by different metrics (see Section 1.7). The resulting algorithm is fast and stable. However, a further increase in the speed may be obtained by using the derivatives given in Section 1.3.[1]

---

[1]We decided not to implement a numerical optimization algorithm using derivatives, because an increase in the speed also for the subclasses would have required the implementation of all derivatives for the subclasses. However, after a decision for a particular subclass of GH distributions one should use the derivatives to increase the speed. For instance, the `hyp` program of Blæsild and Sørensen (1992) for the estimation of hyperbolic distributions uses derivatives.

We have implemented the direct solutions for $\mu$ given in Section 1.3. In contrast to the hyperbolic case, estimation of GH parameters for financial return data converges quite often to limit distributions at the boundary of the parameter space. Hence, it was necessary to adapt the algorithm to the parameter restrictions given above. We have also modified the algorithm to estimate parameters for a fixed $\lambda$.

An important point concerning the speed of the estimation is the selection of starting values. To estimate the generalized hyperbolic distribution for a particular data set, we choose starting values by rescaling a symmetric GH distribution with a reasonable kurtosis, e.g. $\xi \approx 0.7$, such that the empirical variance and the variance of the GH distribution are equal. See Section 2.1 for a detailed description of the rescaling mechanism. We tested this approach with data sets from different origins. In general, this choice of starting values leads to good results. The estimation algorithm is stopped if the relative change in the parameter values $(p_{i,n})_{1 \leq i \leq 5}$ in the $n$-th iteration step is smaller than a given constant, i.e.

$$\sum_{i=1}^{5} \left| \frac{p_{i,n+1} - p_{i,n}}{p_{i,n}} \right| < 10^{-10}. \tag{1.21}$$

Although the computational power increases, it is necessary to find a reasonable tradeoff between the introduction of additional parameters and the possible improvement of the fit. Barndorff-Nielsen and Blæsild (1981) mentioned the flatness of the likelihood function yet for the hyperbolic distribution. The variation in the likelihood function of the GH distribution is even smaller for a wide range of parameters (see Section 1.6 below). Consequently, the generalized hyperbolic distribution applied as a model for financial data leads to overfitting. This will become more obvious in the following sections.

The first four centered moments of return distributions yield simple and useful econometric interpretations: trend, riskiness, asymmetry and the probability of extreme events. For fixed $\lambda$ subclasses with four parameters corresponding to the mentioned characeristics are obtained. Therefore, it seems to be appropriate to model return data with one of the subclasses. However, from a mathematical point of view, GH distributions, as the most general distributions which include all subclasses, are more interesting.

Because of the restrictions on the parameter values and the flatness of the likelihood function, it seems not to be possible to use standard minimization algorithms. These ready implemented routines often assume that the parameters and the value of the function have the same order of magnitude and that the gradient is not too small (Press, Teukolsky, Vetterling, and Flannery 1992). Although we have no theoretically assured convergence of our algorithm, tests with different starting values reveal that for financial data the use of reasonable starting values results in the convergence to a global extremum. In the special case of hyperbolic distributions we estimate the same parameters with our algorithm and the `hyp` program implemented by Blæsild and Sørensen (1992).

The main factor for the speed of the estimation is the number of modified Bessel functions to compute. They are calculated by a numerical approximation described in Press, Teukolsky, Vetterling, and Flannery (1992, pp. 236–252). Note, that for $\lambda = 1$ this function appears only in the norming constant. For a data set with $n$ independent observations we need to evaluate $n$ and $n + 1$ Bessel functions for NIG and GH distributions respectively, whereas only one for $\lambda = 1$. This leads to a striking reduction in the time necessary to calculate the likelihood function in the hyperbolic case (see Table 1.11).

We applied the estimator to log-return[2] data from the German stock market and to New York Stock Exchange (NYSE) indices. We define the return of a price process $(S_t)_{t \geq 0}$ for a time interval $\Delta t$, e.g. one day, as

$$X_t = \log S_t - \log S_{t - \Delta t}. \tag{1.22}$$

Thus, the return during $n$ periods is the sum of the one-period returns. The stock data set consists of daily closing prices from January 1988 to May 1994. We had to correct these quoted prices due to dividend payments. The NYSE indices are reported from January 2, 1990 to November 29, 1996. In Tables 1.13 and 1.14 we present the estimated GH, hyperbolic and NIG distributions, and those obtained with $\lambda = 0, 1/2, -3/2$ for both data sets. The tables present also the value of the log-likelihood function and the second and third parametrizations $(\varrho, \zeta)$ and $(\chi, \xi)$.

For numerical reasons it is useful to find a suitable subclass of the GH distributions, i.e. choose a $\lambda \in 1/2\ \mathbb{Z}$. The estimates of $\lambda$ for the NYSE indices are scattered around zero. Therefore, it is not possible to choose one subclass for all NYSE indices. For the 30 german stocks in the DAX the estimates for $\lambda$ range from $-2.4$ to $0.8$, but for 23 of 30 stocks we obtain $-2 < \lambda < -1.4$. Similar results are obtained for the DAX (Eberlein and Prause 1998).

Moreover, a volatility approach based on the mixing GIG distribution leads to a comparable range of estimates for $\lambda$. Wirth (1998) estimated GIG distributions for historical volatility estimates. Then he compared GH distributions constructed from the estimated GIG distributions (by means of the mixing representation (1.6)) with the empirical distribution of the returns: both distributions are not very distinct.[3] The estimates for $\lambda$ are below $-1/2$. In both approaches, neither NIG nor hyperbolic distributions are inside the range of the estimates for $\lambda$.

In view of the obtained range of $\lambda$, the subclass of the generalized hyperbolic distribution with $\lambda = -3/2$ is interesting. It has the following density, mean and variance

$$\widetilde{h}(x; \alpha, \beta, \delta, \mu) = a(-3/2, \alpha, \beta, \delta) \frac{\mathrm{K}_2 \left( \alpha \sqrt{\delta^2 + (x - \mu)^2} \right)}{\delta^2 + (x - \mu)^2} e^{\beta(x - \mu)} \tag{1.23}$$

$$a(-3/2, \alpha, \beta, \delta) = \frac{\alpha^2 \delta^{5/2}}{\pi (\alpha^2 - \beta^2)^{1/4} (1 + \zeta)} e^{\zeta}$$

$$\mathrm{E}\, X = \mu + \frac{\beta \delta^2}{\delta \sqrt{\alpha^2 - \beta^2} + 1}$$

$$\mathrm{Var}\, X = \delta^2 \left( \frac{1}{\zeta} + \frac{\beta^2}{\alpha^2 - \beta^2} \frac{1}{\zeta + 1} \right) \frac{\zeta}{\zeta + 1}$$

The number of parameters is reduced to four and the Bessel function K is of integer order, which simplifies numerical calculations and the estimation. Unfortunately, this subclass is not closed under convolution and the function contains a Bessel function outside the norming constant $a$. Hence, the estimation is considerably more time-consuming than for hyperbolic distributions. Note, that the variation of the likelihood for the GH distributions and the subclasses is very small. We will investigate this in Section 1.4 in a more detailed way.

---

[2]We use "returns" and "log-returns" synonymously.

[3]Wirth (1998) obtained the following $\hat{\lambda}$ for historical volatility estimators applied to the DAX and defined in Section 2.5: Hist15 $-0.85$, Hist20 $-1.27$, Hist30 $-2.28$. For the VDAX he obtained $\hat{\lambda} = -3.99$, but the GH distribution constructed by this GIG estimate is not close to the distribution of the empirical returns (estimation period: January 1992 to December 1997).

Naturally the GH estimates (in the Tables 1.13 and 1.14) have a higher likelihood than the nested subclasses. For the German stocks we usually obtain values for $\lambda$ below $-3/2$ and hence under all subclasses the likelihood is maximal for $\lambda = -3/2$ and decreases with increasing $\lambda$.

The values of $\chi$ are usually close to zero, which indicates a symmetric distribution. For the subclasses we obtain $\xi$ values from 0.65 to 0.86. Recall that $\chi$ and $\xi$ give the position of the GH estimate in the shape triangle corresponding to the subclass.

For seven German stocks (Allianz-Holding, Bayerische Vereinsbank, Commerzbank, Karstadt, MAN, Mannesmann, Siemens) and the NYSE Composite Index the estimated GH distribution converge to the boundary of the parameter space as $\beta \to \alpha$, $\lambda < 0$, $0 < \delta$. In terms of the other parametrizations this means $\chi, \xi, \varrho \to 1$ and $\zeta \to 0$.

**Lemma 1.24.** *The limit of the generalized hyperbolic distribution for $\beta \to \alpha$ and $\lambda < 0$ has the following density*

$$\bar{h}(x; \lambda, \alpha, \delta, \mu) = \frac{2^{\lambda+1}}{\sqrt{2\pi}\,\Gamma(-\lambda)\delta^{2\lambda}\,\alpha^{\lambda-1/2}}\, \mathrm{K}_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)\exp(\alpha(x-\mu)).$$

*Proof.* For $x \downarrow 0$, $\lambda > 0$ the following property of the modified Bessel function $\mathrm{K}_\lambda$ holds

$$\mathrm{K}_\lambda(x) \sim \Gamma(\lambda)2^{\lambda-1}x^{-\lambda}. \tag{1.25}$$

We also use the fact that $\mathrm{K}_\lambda(x) = K_{-\lambda}(x)$. Hence, the norming factor $a$ given in Definition 1.1 simplifies in the following way

$$
\begin{aligned}
a(\lambda, \alpha, \beta, \delta) \quad &= \quad \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi}\,\alpha^{\lambda-1/2}\,\delta^\lambda\,\mathrm{K}_{-\lambda}\left(\delta\sqrt{\alpha^2 - \beta^2}\right)} \\[2mm]
&\xrightarrow[\alpha\downarrow\beta]{} \quad \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi}\,\alpha^{\lambda-1/2}\,\delta^\lambda\Gamma(-\lambda)2^{-\lambda-1}\delta^\lambda(\alpha^2 - \beta^2)^{\lambda/2}} \\[2mm]
&= \quad \frac{2^{\lambda+1}}{\sqrt{2\pi}\,\Gamma(-\lambda)\delta^{2\lambda}\,\alpha^{\lambda-1/2}}.
\end{aligned}
$$
□

The parametrization in this limiting case is 4-dimensional, but a substantial change appears only in the norming factor.

Since $\beta \to \alpha$ the parameter $\psi = \alpha^2 - \beta^2$ of the GIG distribution in the mixture representation tends to 0. For $\mu = 0$ we obtain then the Student-t distribution as a mixture of normals with inverse gamma distributions (Barndorff-Nielsen and Blæsild 1983). The obtained distribution for $\mu \neq 0$ is not equal to a noncentral t-distribution since the latter one is a mixture only in the variance with a constant mean and not a variance-mean mixture as the GH distribution (Johnson and Kotz 1970).

## 1.3   Derivatives of the Log-Likelihood Function

We obtain the following expressions for the derivatives of the log-likelihood function of the GH distributions, which we denote by $L$. See Appendix B for the definition and properties of the functions $R_\lambda$ and $S_\lambda$. We apply in the in the differentiation the well-known properties of the Bessel function $\mathrm{K}_\lambda$ given in the appendix and the abbreviation $k_\lambda(x) = d\,\mathrm{K}_\lambda(x)/d\lambda$ for

the derivative of the Bessel function with respect to order. We skip the tedious details of the differentiation.

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\lambda}L &= n\left[\frac{1}{2}\ln\frac{\alpha^2-\beta^2}{\alpha\delta} - \frac{k_\lambda(\delta\sqrt{\alpha^2-\beta^2})}{\mathrm{K}_\lambda(\delta\sqrt{\alpha^2-\beta^2})}\right] \\
&\quad + \sum_{i=1}^{n}\left[\frac{1}{2}\ln(\delta^2+(x_i-\mu)^2) + \frac{k_{\lambda-1\!/2}(\alpha\sqrt{\delta^2+(x_i-\mu)^2})}{\mathrm{K}_{\lambda-1\!/2}(\alpha\sqrt{\delta^2+(x_i-\mu)^2})}\right] \\
\frac{\mathrm{d}}{\mathrm{d}\alpha}L &= n\frac{\delta\alpha}{\sqrt{\alpha^2-\beta^2}}R_\lambda(\delta\sqrt{\alpha^2-\beta^2}) \\
&\quad - \sum_{i=1}^{n}\sqrt{\delta^2+(x_i-\mu)^2}R_{\lambda-1\!/2}(a\sqrt{\delta^2+(x_i-\mu)^2}) \\
\frac{\mathrm{d}}{\mathrm{d}\beta}L &= n\left[-\frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}}R_\lambda(\delta\sqrt{\alpha^2-\beta^2}) - \mu\right] + \sum_{i=1}^{n}x_i \\
\frac{\mathrm{d}}{\mathrm{d}\delta}L &= n\left[-\frac{2\lambda}{\delta} + \sqrt{\alpha^2-\beta^2}R_\lambda(\delta\sqrt{\alpha^2-\beta^2})\right] \\
&\quad + \sum_{i=1}^{n}\left[\frac{(2\lambda-1)\delta}{\delta^2+(x_i-\mu)^2} - \frac{\alpha\delta R_\lambda(\alpha\sqrt{\delta^2+(x_i-\mu)^2})}{\sqrt{\delta^2+(x_i-\mu)^2}}\right] \\
\frac{\mathrm{d}}{\mathrm{d}\mu}L &= -n\beta + \sum_{i=1}^{n}\frac{x_i-\mu}{\sqrt{\delta^2+(\mu-x_i)^2}} \\
&\quad \times \left[\frac{2\lambda-1}{\sqrt{\delta^2+(x_i-\mu)^2}} - \alpha R_{\lambda-1\!/2}(\alpha\sqrt{\delta^2+(x_i-\mu)^2})\right].
\end{aligned}
$$

Therefore, we obtain direct solutions of the likelihood equations for $\beta$ and $\mu$. From $\frac{\mathrm{d}}{\mathrm{d}\beta}L=0$ and $\frac{\mathrm{d}}{\mathrm{d}\mu}L=0$ we get

$$
\mu = -\frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}}R_\lambda(\delta\sqrt{\alpha^2-\beta^2}) + \frac{1}{n}\sum_{i=1}^{n}x_i \tag{1.26}
$$

$$
\begin{aligned}
\beta = \frac{1}{n}\sum_{i=1}^{n}&\frac{x_i-\mu}{\sqrt{\delta^2+(x_i-\mu)^2}} \\
&\times\left[\frac{2\lambda-1}{\sqrt{\delta^2+(x_i-\mu)^2}} - \alpha R_{\lambda-1\!/2}(\alpha\sqrt{\delta^2+(x_i-\mu)^2})\right].
\end{aligned} \tag{1.27}
$$

The solutions of the likelihood equations for the other parameters are obtained by maximizing the log-likelihood equation with respect to $(\lambda,\alpha,\delta,\beta)$. The location parameter $\mu$ is then found by (1.26). It is of course preferable to maximize $(\lambda,\log\alpha,\log\delta,\beta)$ to circumvent the positivity condition on $\alpha$ and $\delta$.

11
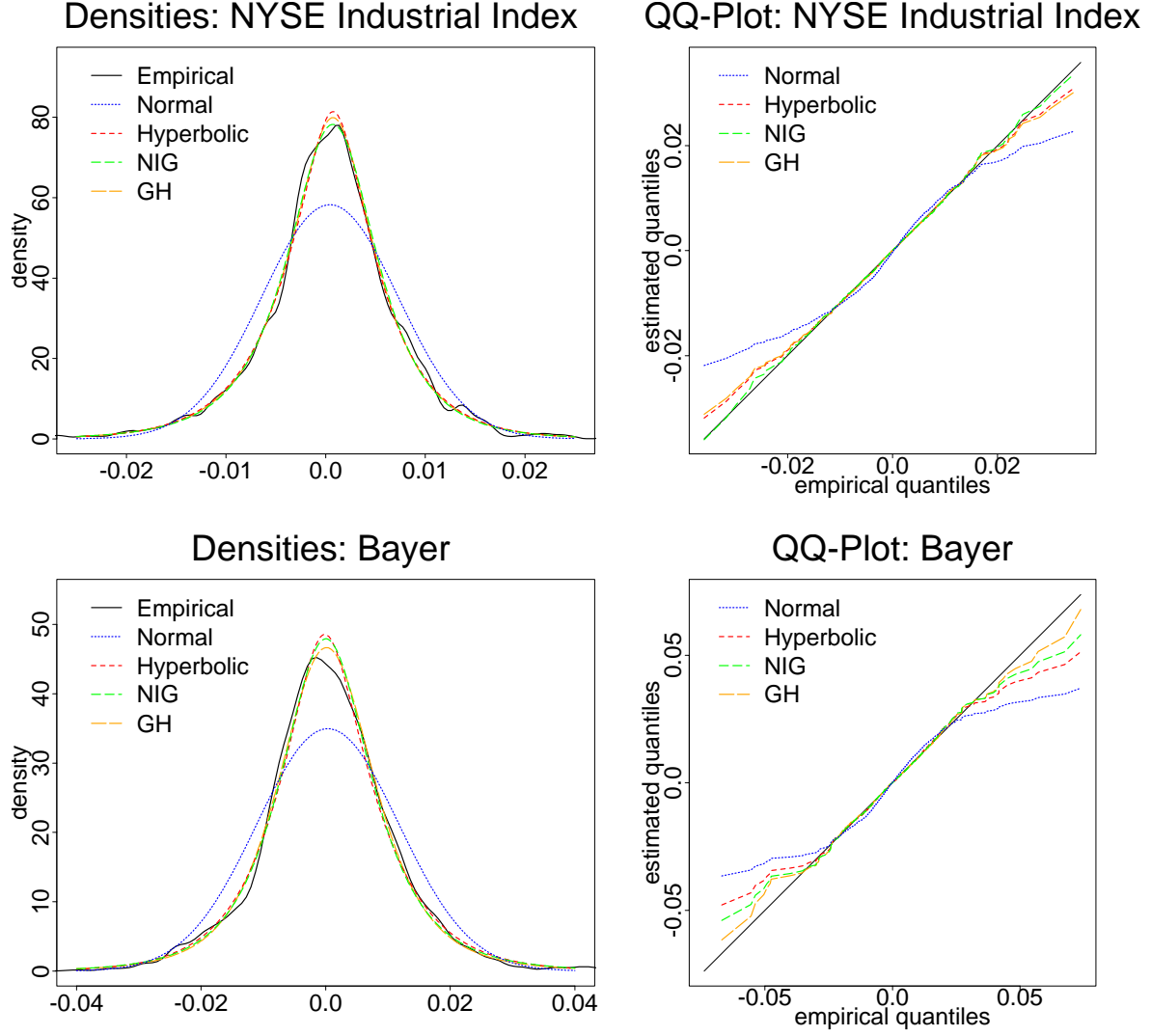
Figure 1.2: *NYSE Industrial Index and Bayer.*

## 1.4  Comparison of the Estimates and Tests

The aim of this section is to analyze and compare the goodness of fit of the generalized hyperbolic distributions and their subclasses. For a first graphical comparison we provide plots of the densities, and qq-plots for the NYSE Industrial Index and Bayer share in Figure 1.2. Both types of plot reveal that generalized hyperbolic distributions provide excellent fits to empirically observed return distributions. It is obvious that generalized hyperbolic distributions are leptokurtic, i.e. the peak in the center is higher and there is more mass in the tails than for the normal distribution.

We also compare the GH estimates with fitted normal distributions statistically. As a measure for the goodness of the fit we use various distances between the fitted and the empirical cumulative density function (CDF): The Kolmogorov distance is defined as the supremum over the absolute differences between two cumulative density functions. We also

compute $L^1$ and $L^2$ distances of the CDFs. The Anderson & Darling statistic is given by

$$\text{AD} = \max_{x \in \mathbb{R}} \frac{|F_{\text{emp}}(x) - F_{\text{est}}(x)|}{\sqrt{F_{\text{est}}(x)(1 - F_{\text{est}}(x))}}, \tag{1.28}$$

where $F_{\text{emp}}$ and $F_{\text{est}}$ are the empirical and the estimated CDFs. We apply this statistic because it pays more attention to the tails of the distribution (see also Hurst, Platen, and Rachev 1995) and therefore hints at the possibility to model the probability of extreme events with a given distribution. In Table 1.3 we give the results for some share values included in the German stock index DAX.

Table 1.3: *Goodness of fit of the GH, NIG, hyperbolic and normal distribution. Different metrics are applied to measure the difference between the estimated and the empirical cumulative density functions.*

|  | GH | NIG | HYP | normal |
|---|---|---|---|---|
| Kolmogorov distance | | | | |
| Allianz-Holding | 0.0016 | 0.0018 | 0.0019 | 0.0097 |
| Bayer | 0.0164 | 0.0167 | 0.0160 | 0.0593 |
| Daimler Benz | 0.0122 | 0.0122 | 0.0120 | 0.0628 |
| Anderson & Darling statistic | | | | |
| Allianz-Holding | 0.1301 | 0.5426 | 3.0254 | 5.84$e$07 |
| Bayer | 0.0604 | 0.0884 | 0.1462 | 17.8506 |
| Daimler Benz | 0.1094 | 0.5533 | 4.0783 | 6.58$e$09 |
| $L^1$-distance of CDFs | | | | |
| Allianz-Holding | 0.0004 | 0.0006 | 0.0007 | 0.0024 |
| Bayer | 0.0003 | 0.0003 | 0.0004 | 0.0015 |
| Daimler Benz | 0.0004 | 0.0005 | 0.0005 | 0.0021 |
| $L^2$-distance of CDFs | | | | |
| Allianz-Holding | 0.0016 | 0.0018 | 0.0019 | 0.0097 |
| Bayer | 0.0011 | 0.0012 | 0.0015 | 0.0070 |
| Daimler Benz | 0.0014 | 0.0013 | 0.0014 | 0.0085 |

For all analyzed metrics we get better results for the GH distributions and their subclasses than for the normal distribution. The poor fit of normal distributions to the heavier tails of returns comes to light in the values of the Anderson & Darling statistic. Looking at the statistics for the GH, NIG and HYP distributions, no striking differences are found. Recalling the flatness of the likelihood function and the proximity of the log-likelihood values in Tables 1.13 and 1.14, this result is no surprise and underlines the overfitting of generalized hyperbolic distributions with arbitrary $\lambda$. The values of the Kolmogorov and $L^2$ distances of GH, NIG and hyperbolic distributions are very close, and the distribution with the highest value changes. Both, Anderson & Darling statistic and the $L^1$ distance reveal the following ranking in the fit: GH, NIG, hyperbolic and normal distribution.

Another way to analyze the discrepance of GH distributions to their subclasses is to perform a statistical test under the hypothesis that the true distribution is included in the

Table 1.4: *Likelihood ratio test of GH estimates under the hypothesis of a hyperbolic resp. NIG distribution.*

| | Hyperbolic | | NIG | |
| | statistic | p-value | statistic | p-value |
|---|---|---|---|---|
| NYSE indices, 1746 observations, January 2, 1990 to November 11, 1996 | | | | |
| Composite | 0.097 | 0.756 | 3.003 | 0.083 |
| Finance | 0.758 | 0.384 | 0.234 | 0.629 |
| Industrial | 1.245 | 0.264 | 0.915 | 0.339 |
| Transport | 15.1 | 0.0001 | 7.107 | 0.008 |
| Utility | 1.614 | 0.204 | 0.16 | 0.689 |
| German stock, 1598 observations, January 1, 1988 to May 24, 1994 | | | | |
| BASF | 21.435 | $3.6e{-}06$ | 8.397 | 0.004 |
| Bayer | 15.375 | 0 | 4.3 | 0.038 |
| Daimler Benz | 18.385 | $1.8e{-}05$ | 4.476 | 0.034 |
| Deutsche Bank | 11.604 | 0.001 | 0.764 | 0.382 |
| Siemens | 32.195 | $1.3e{-}08$ | 12.113 | 0.001 |
| RWE | 18.196 | $1.9e{-}05$ | 0.005 | 0.946 |

hyperbolic resp. NIG subclass of GH distributions. In Table 1.4 we provide the results of a likelihood ratio test.

Looking at the p-values, we see that the NIG and the hyperbolic hypothesis is not rejected at a 5% niveau for all NYSE indices but the NYSE Transport Index. Both hypotheses are rejected in most cases for the German stocks. This is no surprice for a large data set, since the likelihood ratio test is consistent (see Theorem 6.60 and the following remarks in Witting and Müller-Funk 1995). Nevertheless, this result hints at an even better fit of NIG and hyperbolic distributions for the NYSE indices in comparison to German stocks. This is also visible in the qq-plots on page 12.

## 1.5 High-Frequency Data

Generalized hyperbolic distributions are also appropriate for the modelling of intraday data. We explore a data set published by Olsen & Associates (1998) in preparation of the HFDF II conference 1998. See Appendix A.1 for a detailed description of the data set. We estimate distributions of intraday returns for different time lags, using data recorded from the global foreign exchange (FX) market and from the US-market. For high-frequency data we follow Guillaume et al. (1997) in the definition of a log-price

$$p(t) = [\log p_{\mathrm{ask}}(t) + \log p_{\mathrm{bid}}(t)]/2 \qquad (1.29)$$

and the corresponding return

$$r(t) = p(t) - p(t - \Delta t), \qquad (1.30)$$

where $\Delta t$ is the time interval. We estimate the GH parameters after removing all zero-returns. Although this is only a provisional approach to focus on time periods where trading
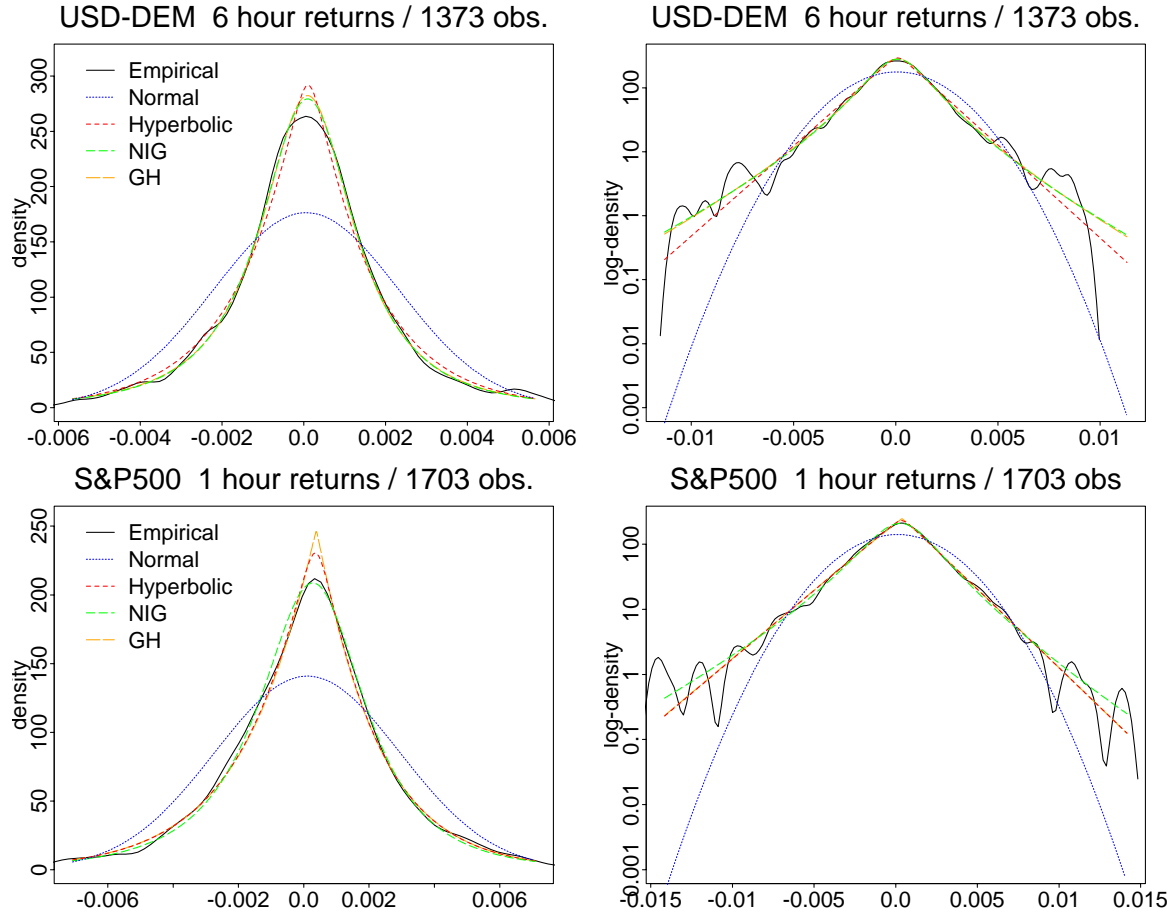
Figure 1.5: *Densities and log-densities of high frequency data.*

takes place, we can conclude that under the assumption of independence the class of GH distributions provides excellent fits to the empirically observed distributions. This assumption is of course not realistic, considering the seasonalities and long-range dependence usually observed in these time series. The microstructure one finds in these globally traded assets needs much more attendance, particularly in comparison to daily data from a single market. However, the results given in Figure 1.5 encourage further research. In particular, stochastic volatility models with generalized hyperbolic marginals may be interesting.

## 1.6 Finite-Sample Properties

In this section we investigate the stability of the estimation of GH distributions by a simulation study. We generate random numbers from the GH distribution estimated for the Bayer returns (see Table 1.14) using the quantile function and a random number generator on $[0, 1]$, and we produce data sets with different sample sizes. Note, that the choice of the sampling distributions restricts the validity of the following results to typical log-return distributions of financial assets. On page 17 we show the results of the simulation. Similar results are obtained for other sampling distributions.

**Volatility and Quantiles**

First, we assume that the true distribution is GH, compute its theoretical standard deviation and 1% quantile, and compare both with the values resulting from likelihood estimates of hyperbolic and NIG distributions and the nonparametric estimates. The results are given in Table 1.6.

The first observation is, that the estimation of the standard deviation is strikingly poor. For 250 observations, which corresponds to the number of trading days in one year, the mean absolute errors are larger than 5% and for 1598 ($\approx 6.5$ years) about 3%. In most cases we observe an underestimation of the volatility and the nonparametric estimate of the volatility gives better results.

On the contrary, the parametric results for the estimation of 1% quantiles are better than the non-parametric. However, the mean absolute errors measured in percent are even larger than the corresponding errors for the standard deviation.

**Skewness, Kurtosis, and Subclass**

Due to the flat log-likelihood the comparison of single parameter values of the sampling and the estimated distribution does not make sense. Therefore, we prefer to compute the error in the estimate for the subclass $\lambda$, the Kolmogorov distance $\max_{x \in \mathbb{R}} |F_{\text{sampling}}(x) - F_{\text{est}}(x)|$ and the Anderson & Darling statistic defined in (1.28). The results are given in Table 1.7.

Latter Table shows that only for large sample sizes the parameter $\lambda$ is close to the sampling distribution. This reveals that the estimation of the subclass, characterized by $\lambda$, is difficult. For instance, the mean absolute error in the estimation of $\lambda$ for 1598 observations is of the same order of magnitude as the difference in $\lambda$ between NIG and hyperbolic distributions. Recall from Section 1.4 that the difference between the subclasses in terms of the likelihood is small. Of course, it is not possible to find financial time series at any given length without getting in trouble with changes of regime.

Table hints at the number of observations necessary for a desired precision of the estimate. The Kolmogorov and A&D distance decrease with increasing sample size. From these results we conclude—as a rule of thumb—that at least 250 observations are necessary to obtain a reasonable fit.

Another way to visualize the stability of the GH estimates with respect to the shape of the distribution is given in Figure 1.8. We generated 100 i.i.d. random samples of length 250 and 1598 for the hyperbolic distribution estimated for the Bayer stock. The shape of the hyperbolic distributions estimated from these random samples is shown in the shape triangles. Obviously, it is preferable to estimate shape parameters from time series longer than one year.

## 1.7 Minimum-Distance Estimation

In this section we apply different estimation methods by replacing the log likelihood function with other score functions. We want to know whether one of these different approaches yields a better fit, in particular, to tails of GH distributions. This may improve the modelling of the probability of extreme events. We estimate parameters for the GH, NIG and hyperbolic distributions using the metrics already applied in Section 1.4. The results are given on page 33 for Deutsche Bank returns. Note, that the numerical differences in the values for GH, NIG and hyperbolic distributions are small for each distance. For the minimal Kolmogorov

Table 1.6: *Simulation study: estimation of standard deviation and quantiles. Sampling distribution: Bayer ML estimate for the GH distribution with standard deviation 0.01147 and 1%-quantile −0.0298; 1000 iterations.*

| Sample size | 250 random numbers | | | | 1598 random numbers | | |
|---|---|---|---|---|---|---|---|
| Average | Error % | AbsErr % | SquErr | | Error % | AbsErr % | SquErr |
| *Standard Deviation* | | | | | | | |
| Empirical | 0.21 | 7.07 | $1.15e{-}6$ | | −0.26 | 2.95 | $1.78e-7$ |
| Hyperbolic | −2.95 | 5.92 | $7.13e{-}7$ | | −3.20 | 3.51 | $2.29e-7$ |
| NIG | −1.05 | 6.40 | $8.69e{-}7$ | | −1.63 | 2.84 | $1.62e-7$ |
| *1%-Quantile* | | | | | | | |
| Empirical | 3.67 | 14.33 | $2.80e{-}5$ | | 0.79 | 5.94 | $5.03e-6$ |
| Hyperbolic | 3.39 | 9.70 | $1.27e{-}5$ | | 2.92 | 4.40 | $2.66e-6$ |
| NIG | −0.08 | 11.38 | $1.86e{-}5$ | | −0.27 | 4.53 | $2.99e-6$ |

Table 1.7: *Stability of the estimation of the full GH distribution, i.e. with arbitrary $\lambda$. Sampling distribution: GH maximum-likelihood estimate for Bayer; 100 iterations.*

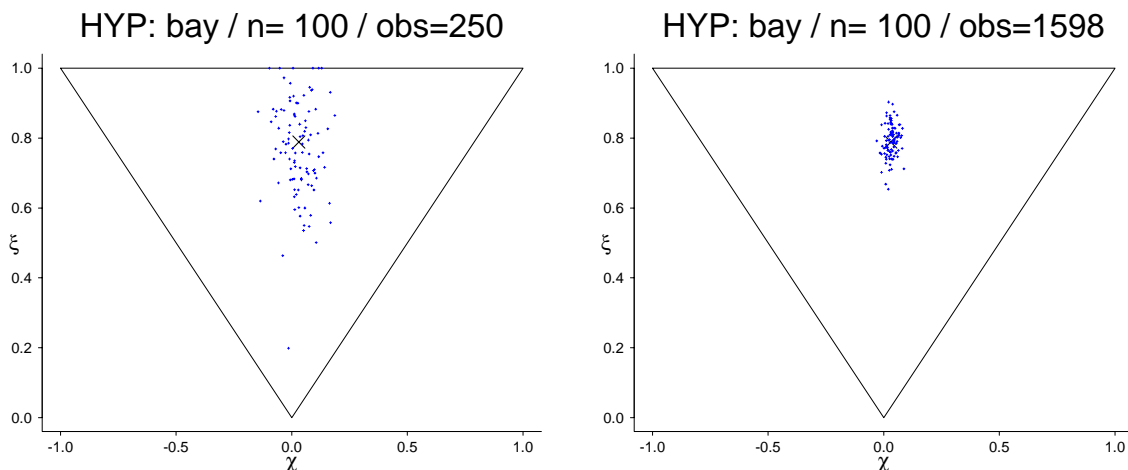| Sample Size | Mean Absolute Errors | | Mean | Mean |
|---|---|---|---|---|
| | $\lambda$ | St Dev | Kolm. dist. | A&D-stat. |
| 50 | 9.182 | 0.00161 | 0.0937 | 0.908 |
| 100 | 4.31 | 0.00128 | 0.0586 | 0.566 |
| 150 | 3.43 | 0.00105 | 0.0451 | 0.433 |
| 200 | 3.572 | 0.00268 | 0.0419 | 0.399 |
| 250 | 2.405 | 0.00356 | 0.035 | 0.333 |
| 500 | 2.127 | 0.000561 | 0.0262 | 0.25 |
| 750 | 1.874 | 0.000549 | 0.0233 | 0.221 |
| 1000 | 1.81 | 0.00104 | 0.0189 | 0.178 |
| 1598 | 1.33 | 0.000362 | 0.0153 | 0.145 |
| 2500 | 0.976 | 0.000351 | 0.0119 | 0.112 |
| 5000 | 0.484 | 0.000205 | 0.00721 | 0.068 |

Figure 1.8: *Stability of GH estimates: Estimation of a hyperbolic distribution from 250 resp. 1598 random variates. This procedure is repeated 100 times.*

distance and the minimal Anderson & Darling distance we observe values for $\lambda$ close to $-1/2$, which indicates that NIG distributions provide a very good fit with respect to these metrics.

Is it useful to use different metrics for the estimation of distributions from log-returns in Finance? To answer this question, we compare the empirical skewness and kurtosis with those values for the estimated distributions. The exact values of the skewness and kurtosis for a specified generalized hyperbolic distribution are computed by the formulas given in Barndorff-Nielsen and Blæsild (1981). Both values are complicated expressions in terms of modified Bessel functions. The results are given in Table 1.15. They indicate that the Anderson & Darling statistic and the Kolmogorov distance have poor finite-sample properties as estimators, in particular, in comparison to $L^p$ distances of CDFs or the maximum likelihood approach. On the one hand, the kurtosis of estimated generalized hyperbolic distributions is always closer to the empirical kurtosis. On the other hand estimated generalized hyperbolic distributions are sometimes skewed in the other direction than the empirical distribution. Similar results are obtained for other stock data sets. In general, the Anderson & Darling statistic and the Kolmogorov distance yield estimates for which skewness and kurtosis deviate in an irregular pattern from the empirical values. The estimates based on $L^p$-norms are closer to the empirical kurtosis, but the estimation of the skewness is rather poor. We obtain the best fits to the empirial skewness and kurtosis, also for other datasets, following the maximum likelihood approach. Therefore, considering the finite-sample properties, it is not favourable, to replace the ML approach by one of the proposed distances.

## 1.8 Measures of Risk

A good fit of the empirically observed semi-heavy tails is also important for the estimation of *Value-at-Risk* (VaR). The motivation to introduce the concept of Value-at-Risk was the necessity to quantify the risk for complex portfolios in a simple way. The VaR to a given *level of probability* $\alpha$ is defined as the maximal loss inherent to a portfolio position over a future *holding period* which is exceeded only with a probability of $\alpha \in (0,1)$. This level of probability is typically chosen as 1% or 5%. It should not be confused with the confidence
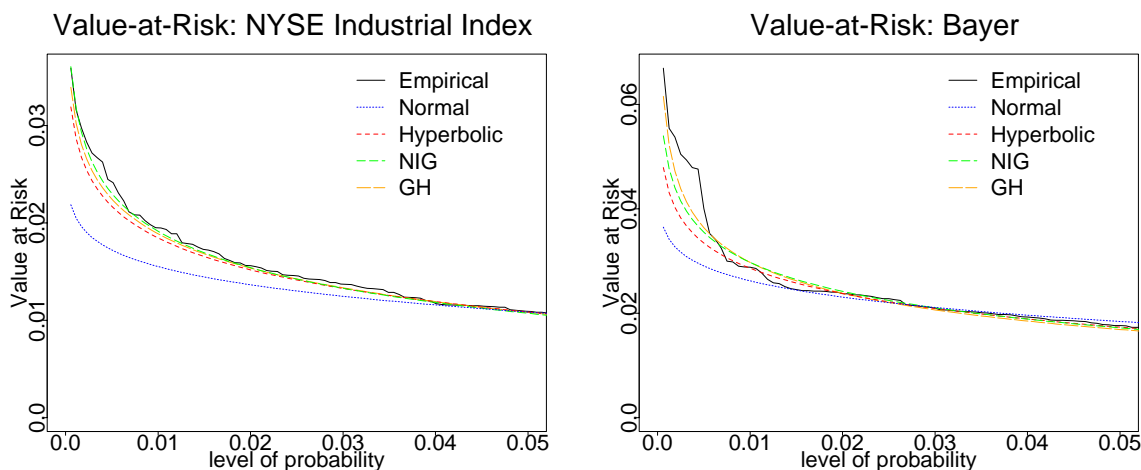
Figure 1.9: *VaR of a portfolio with linear risk and the value of one currency unit (US-$ or Deutsche Mark). The exposure period is one trading day. We compare the empirical VaR at different levels of probability to the estimated VaR using GH, NIG, hyperbolic and normal distributions.*

level. We look at the whole interval of levels of probability. This approach corresponds to the multivariate approach in Davé and Stahl (1997). They propose to measure VaR for a whole range of levels of probability $\alpha$. Since we are interested in extreme events, we have only a small number of observations; e.g. in a 99% quantile we obtain about 16 observations for our data sets. Looking into VaR as a function of $\alpha$ gives an impression for the behaviour of a risk measure.

First of all, we analyze VaR for portfolios with linear risk, e.g. portfolios consisting of only one stock or index. The VaR estimates for the GH, NIG and hyperbolic distribution are given in Figure 1.9. Obviously the class of generalized hyperbolic distributions and its subclasses provide better fits to the empirical VaR—in particular for small levels of probability—than the normal distribution.

A detailed discussion of the various approaches to risk measurement especially for multivariate price processes is given in Chapter 5. The good description of the risk inherent in a financial asset by GH distributions may also be used as a basic ingrediente to portfolio optimiziation approaches. For instance, Sulem (1999) has shown that a smaller part of the capital is invested in the risky asset, if the process of the risky asset is non-Gaussian.

## 1.9 Subordination, Lévy-Khintchine Representation

In the following section we derive the Lévy-Khintchine representation of GH distributions for arbitrary $\lambda$. In the case $\lambda = 1$ and $\lambda = -1/2$ the corresponding densities have been derived by Keller (1997, Proposition 51) and Barndorff-Nielsen (1998). However, the Lévy-Khintchine representation for arbitrary $\lambda$ is rather unpleasant, therefore we add subsequently a simpler version for positive $\lambda$.

Apart from the fact that the subordination allows to calculate the Lévy-Khintchine representation, it gives also a link to other approaches to explain the behaviour of asset prices by a simpler process viewed on *business time scales*. See Schnidrig and Würtz (1995) for

empirically constructed business times scales for high-frequency data, and Geman, Madan, and Yor (1998) for time-transformed Brownian motions as a model for asset prices.

## Subordination

A straightforward approach to calculate the Lévy density of the GH distribution is based on the representation of the GH Lévy motion as a Brownian motion $W(t)$ with independent random time $Y(t)$. This representation is denoted as *subordination* (Huff 1969).

**Proposition 1.31 (Lévy-Khintchine Representation).** *A variable $X$ following a centered and symmetric GH with parameters $(\lambda, \alpha, 0, \delta, 0)$, $\lambda \in \mathbb{R}$ has the Lévy-Khintchine representation of the characteristic function*

$$\widehat{X}(u) = f_{W(Y(1))} = \exp\left(\int_\infty^\infty (\cos(ux) - 1)dM_X(x)\right) \tag{1.32}$$

*with spectral measure*

$$M_X(x) = -\text{sign}(x) \int_0^\infty \int_{-\infty}^{-|x|} \frac{1}{\sqrt{2\pi t}} \exp(-y^2/2t)dy\, dM_Y(t). \tag{1.33}$$

*The spectral measure $M_Y$ of the random time $Y$ has the density*

$$h(x) = \left[\delta^2 \int_0^\infty e^{-x\xi} g_\lambda(2\delta^2\xi)d\xi + \max\{0, \lambda\}\frac{\lambda}{x}\right] e^{-\alpha^2 x/2}, \tag{1.34}$$

*with*

$$g_\lambda(x) = \left[(\pi^2/2)x\left(J_{|\lambda|}^2(\sqrt{x}) + N_{|\lambda|}^2(\sqrt{x})\right)\right]^{-1}. \tag{1.35}$$

*Further, $J$ and $N$ are modified Bessel functions.*[4]

*Proof.* As a consequence of the variance-mean mixture, GH Lévy motions may be written as subordinated processes $W(Y(T))$ in the sense of Huff (1969). Here $W(t)$ denotes a Brownian motion and $Y(t)$ an independent random time with characteristic function

$$f_{Y(T)}(u) = \exp\left(T\left(iu\gamma_Y + \int_0^\infty (e^{iux} - 1)dM_Y(x)\right)\right).$$

From Theorem 4 in Huff (1969) follows that the subordinated process $W(Y(T))$ has the following Lévy triplet $\gamma = 0$, $\sigma^2 = T\gamma_x$ and

$$M(x) = TM_0(x) = \begin{cases} T\int_0^\infty \int_{-\infty}^x \frac{1}{\sqrt{2\pi t}} \exp(-y^2/2t)dy\, dM_Y(t), & x < 0 \\ -TM_0(-x), & x > 0. \end{cases}$$

Consequently one obtains for the characteristic function at $T = 1$ the following expression

$$f_{W(Y(1))} = \exp\left(-\gamma_Y u^2/2 + \int_{-\infty}^\infty (\cos(ux) - 1)dM_0(x)\right).$$

Replacing $M_Y$ with the spectral measure of GIG distributions yields the result above. The density of $M_Y$ is given in Barndorff-Nielsen and Shephard (1998, Theorem 4.1).[5] □

---

[4]Sometimes $N$ is denoted by $Y$, see e.g. Halgreen (1979)

[5]See Table 1.1 for an overview over the various parametrizations.

**Remark 1.36.** *The Lévy triplet of GH symmetric and centered GH distributions contains no Brownian part. Of course, assuming a drift $\mu \neq 0$ does not change this.*

*Moreover, the Lévy-Khintchine representation of skewed GH distributions are related to symmetric GH distributions Lemma 1.51. Obviously, also skewed GH Lévy motions are pure jump processes.*

## Positive $\lambda$

In the case $\lambda > 0$ we get a simpler Lévy-Khintchine representation which does not translate to negative $\lambda$, since a integral representation similar to Theorem B.19 is not valid.

**Lemma 1.37.** *Denote by $\xi(u) = \int_0^\infty e^{-uz} \operatorname{gig}(z)dz$ the Laplace transform of the GIG distribution. Then we have $\varphi_0(u) = \xi\big(\frac{u^2}{2} - i\beta u\big)$ where $\varphi_0$ is the characteristic function of the GH distribution with $\mu = 0$.*

*Proof.* The GH distribution has the following mixture representation $\operatorname{gh}(dx) = \int_0^\infty N(\beta w, w)(dx) \operatorname{gig}(w)dw$ where N is the normal distribution. With Lukasz (1970, Theorem 12.1.1) follows

$$
\begin{aligned}
\varphi_0(u) &= \int e^{iux} \operatorname{gh}(dx) = \int_0^\infty \int e^{iux} N(\beta w, w)(dx) \operatorname{gig}(w)dw \\
&= \int_0^\infty \exp\left(-\left(\frac{u^2}{2} - i\beta u\right)w\right) \operatorname{gig}(w)dw \\
&= \xi\left(\frac{u^2}{2} - i\beta u\right).
\end{aligned}
$$

$\square$

**Lemma 1.38.** *For the symmetric centered case $\beta = \mu = 0$ we obtain the following representation for the Laplace transform of the GIG distribution*

$$
\xi(u) = \left(\frac{\alpha^2}{\alpha^2 + 2u}\right)^{\lambda/2} \frac{K_\lambda\big(\delta\sqrt{\alpha^2 + 2u}\,\big)}{K_\lambda\big(\alpha\delta\big)}.
$$

*Proof.* Note that $\xi(u) = \varphi\big(\sqrt{2u}\,\big)$ holds for $\beta = 0$. The result follows with Lemma 1.37. $\square$

**Lemma 1.39.** *The Laplace transform of the $\operatorname{GIG}(\lambda, \delta^2, \alpha^2 - \beta^2)$ distribution is given by*

$$
\xi(u) = \left(\frac{\alpha^2 - \beta^2}{\alpha^2 - \beta^2 + 2u}\right)^{\lambda/2} \frac{K_\lambda\big(\delta\sqrt{\alpha^2 - \beta^2 + 2u}\,\big)}{K_\lambda\big(\delta\sqrt{\alpha^2 - \beta^2}\,\big)}.
$$

*Proof.* We abbreviate the norming constant of the generalized inverse Gaussian distribution by

$$
b(\lambda, \chi, \psi) := \frac{(\psi/\chi)^{\lambda/2}}{2\,K(\sqrt{\psi\chi})},
$$

21

where $\chi = \delta^2$ and $\psi = \alpha^2 - \beta^2$. Hence, the Laplace transform is given by

$$
\begin{aligned}
\xi(u) &= b(\lambda, \chi, \psi) \int_0^\infty e^{-ux} \exp\left(-\frac{1}{2}\left(\frac{\chi}{x} + \psi x\right)\right) dx \\
&= \frac{b(\lambda, \chi, \psi)}{b(\lambda, \chi, \psi + 2u)} \\
&= \frac{(\psi/\chi)^{\lambda/2}}{2\,\mathrm{K}_\lambda(\sqrt{\chi(\psi + 2u)})} \frac{2\,\mathrm{K}_\lambda(\sqrt{\chi\psi})}{\left(\frac{\psi + 2u}{\chi}\right)^{\lambda/2}} \\
&= \left(\frac{\psi}{\psi + 2u}\right)^{\lambda/2} \frac{\mathrm{K}_\lambda(\sqrt{\chi(\psi + 2u)})}{\mathrm{K}_\lambda(\sqrt{\chi\psi})}.
\end{aligned}
$$

Replacing $\psi$ and $\chi$ completes the proof. $\qquad\square$

For the proof of the following theorem we need some integral representations.

**Lemma 1.40.**

$$
-\ln\left(1 + \frac{u^2}{2z}\right) = \int \frac{e^{iux} - 1 - iux}{|x|} e^{-\sqrt{2z}\,|x|} dx \tag{1.41}
$$

$$
-\ln\left(1 + \frac{u^2}{\alpha^2}\right) = \int \frac{e^{iux} - 1 - iux}{|x|} e^{-\alpha|x|} dx \tag{1.42}
$$

*Proof.* See Keller (1997, Lemma 50c and p. 83) $\qquad\square$

**Theorem 1.43.** *The Lévy-Khintchine representation of the characteristic function of the generalized hyperbolic distribution for $\lambda \geq 0$ is*

$$
\ln\varphi(u) = iu\mu + \int \left(e^{iux} - 1 - iux\right) g(x) dx
$$

$$
g(x) = \frac{e^{\beta x}}{|x|} \left( \int_0^\infty \frac{\exp\left(-\sqrt{2y + \alpha^2}|x|\right)}{\pi^2 y\left(J_\lambda^2(\delta\sqrt{2y}) + Y_\lambda^2(\delta\sqrt{2y})\right)} dy + \lambda e^{-\alpha|x|} \right)
$$

*Proof.* First of all, we assume that $\mu = \beta = 0$ and define $w_u := \delta\sqrt{\alpha^2 + 2u}$. Hence, we can follow

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}u} \ln\xi(u) &= \frac{\mathrm{d}}{\mathrm{d}u} \ln\mathrm{K}_\lambda(w_u) - \frac{\mathrm{d}}{\mathrm{d}u} \ln(\alpha^2 + 2u)^{\lambda/2} \\
&= \frac{\mathrm{K}_\lambda'(w_u)\delta^2/w_u}{\mathrm{K}_\lambda(w_u)} - \frac{\lambda}{2}\frac{2}{\alpha^2 + 2u} \\
&= \frac{\delta^2}{w_u\,\mathrm{K}_\lambda(w_u)}\left(-\frac{\lambda}{w_u}\mathrm{K}_\lambda(w_u) - \mathrm{K}_{\lambda-1}(w_u)\right) - \frac{\lambda\delta^2}{w_u^2} \\
&= -\frac{\delta^2\lambda}{w_u^2} - \frac{\delta^2\,\mathrm{K}_{\lambda-1}(w_u)}{w_u\,\mathrm{K}_\lambda(w_u)} - \frac{\lambda\delta^2}{w_u^2} \\
&= -\frac{2\delta^2\lambda}{w_u^2} - \frac{\delta^2\,\mathrm{K}_{\lambda-1}(w_u)}{w_u\,\mathrm{K}_\lambda(w_u)}
\end{aligned}
$$

22

Since $\lambda \geq 0$, we can use Theorem B.19 to derive the following by integrating the latter identity

$$
\begin{aligned}
\ln \xi(u) &= \int_0^u \left( -\delta^2 \frac{\mathrm{K}_{\lambda-1}(w_v)}{w_v \,\mathrm{K}_{\lambda-1}(w_v)} - \frac{2\delta^2 \lambda}{w_v^2} \right) dv \\
&= -\delta^2 \int_0^\infty \int_0^u \frac{g_\lambda(y)}{y + w_v^2} dv\, dy - \lambda \int \frac{2}{\alpha^2 + 2v} dv \\
&= -\delta^2 \int_0^\infty \int_0^u \frac{dv}{y + \delta^2(\alpha^2 + 2v)} g_\lambda(y) dy - \lambda \ln\left( 1 + \frac{2u}{\alpha^2} \right);
\end{aligned}
$$

here we use that $\int_0^u \frac{1}{C+u} du = \ln(C+u) - \ln(C) = \ln(1 + u/C)$. Denote $\psi = \alpha^2$ and by the change of variable $y \longmapsto \ln(z) = 2\delta^2 z - \delta^2 \psi$ we get

$$
\begin{aligned}
\ln \xi(u) &= -\delta^2 \int_{\psi/2}^\infty \int_0^u \frac{dv}{z+v} g_\lambda\left( 2\delta^2\left(z - \frac{\psi}{2}\right) \right) dz - \lambda \ln\left( 1 + \frac{2u}{\psi} \right) \\
&= -\delta^2 \int_{\psi/2}^\infty \ln\left( 1 + \frac{u}{z} \right) g_\lambda\left( 2\delta^2\left(z - \frac{\psi}{2}\right) \right) dz - \lambda \ln\left( 1 + \frac{2u}{\psi} \right).
\end{aligned}
$$

Hence, with Lemma 1.37 we obtain

$$
\begin{aligned}
\ln \varphi(u) &= \ln \xi\left( \frac{u^2}{2} \right) \\
&= -\delta^2 \int_{\psi/2}^\infty g_\lambda\left( 2\delta^2\left(z - \frac{\psi}{2}\right) \right) \ln\left( 1 + \frac{u^2}{2z} \right) dz - \lambda \ln\left( 1 + \frac{u^2}{\psi} \right).
\end{aligned}
$$

Applying the integral representations of Lemma 1.40 and the change of variable $z \longmapsto \lambda(y) = y + \frac{\psi}{2}$, and interchanging integrals, we obtain

$$
\begin{aligned}
\ln \varphi(u) &= \delta^2 \int_{\psi/2}^\infty g_\lambda\left( 2\delta^2\left(z - \frac{\psi}{2}\right) \right) \int \frac{e^{iux} - 1 - iux}{|x|} e^{-\sqrt{2z}|x|} dx\, dz \\
&\quad + \lambda \int \frac{e^{iux} - 1 - iux}{|x|} e^{-\alpha|x|} dx \\
&= \int \frac{e^{iux} - 1 - iux}{|x|} \left( \int_0^\infty \delta^2 g_\lambda(2\delta^2 y) e^{-\sqrt{2y+\alpha^2}|x|} dy + \lambda e^{-\alpha|x|} \right) dx.
\end{aligned}
$$

Therefore, the Lévy Khintchine representation of the characteristic function in the case of symmetric centered GH distributions is given by

$$
\begin{aligned}
\ln \varphi(u) &= \int (e^{iux} - 1 - iux) g(x) dx \\
g(x) &= \frac{1}{|x|} \left( \int_0^\infty \frac{\exp(-\sqrt{2y+\alpha^2}|x|)}{\pi^2 y (J_\lambda^2(\delta\sqrt{2y}) + Y_\lambda^2(\delta\sqrt{2y}))} dy + \lambda e^{-\alpha|x|} \right),
\end{aligned}
$$

where $\lambda \geq 0$. With Lemma 1.51, we obtain the Lévy measure of the skewed GH distribution. Introducing also $\mu$ proves the desired result. $\qquad\square$
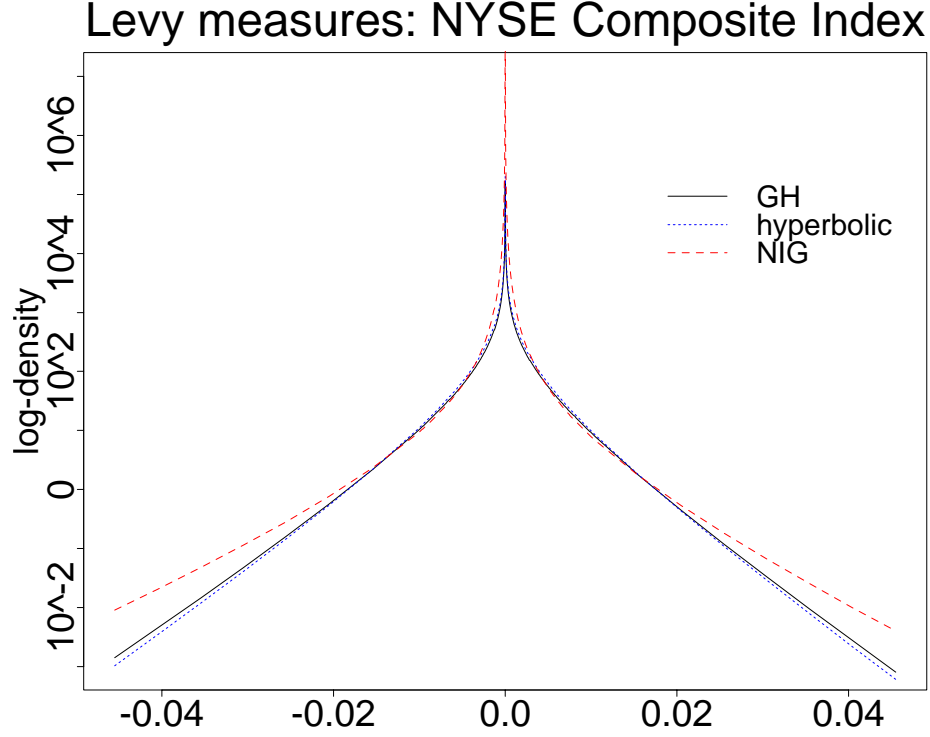
Figure 1.10: *Lévy measure of GH, NIG and hyperbolic distributions.*

In Figure 1.10 we show the Lévy densities of GH distributions estimated for the NYSE composite index. Since the estimated value of $\lambda = 0.84$ is closer to the hyperbolic distribution, in this case the NIG distribution has more large jumps than the GH distribution.

## 1.10 The Generalized Hyperbolic Lévy Motion

Barndorff-Nielsen and Halgreen (1977) showed that the GIG distribution given in Definition 1.7 is infinitely divisible. A direct consequence of this fact is that a mixture of normal distributions with GIG distributions is again an infinitely divisible distribution. Therefore, we can construct Lévy processes based on these distributions. Unfortunately only in the special case of normal inverse Gaussian distributions, i.e. for $\lambda = -1/2$, the family of distributions $\{\text{GH}(-1/2, \alpha, \beta, t\delta, t\mu), \ t \in (0, \infty)\}$ forms a semi-group under convolution. Consequently, the convolutions of generalized hyperbolic distributions in general are no longer generalized hyperbolic. Nevertheless, with the following Lemma we can reduce the calculation of the $n$-fold convolution to the symmetric case.

**Lemma 1.44.** *The convolution semigroup of GH Lévy motions $X$ is characterized by the densities*

$$\text{gh}^{*t}(x; \lambda, \alpha, \beta, \delta, \mu) = \frac{e^{\beta x}}{M_0^t(\beta)} \, \text{gh}^{*t}(x; \lambda, \alpha, 0, \delta, \mu)$$

24

*Proof.* Recall that $a = a(\lambda, \alpha, \beta, \delta)$ was defined as the norming factor of the GH distribution. For a two-fold convolution we obtain:

$$
\begin{aligned}
\mathrm{gh}^{*2}(x) &= \mathrm{gh}^{*2}(x; \lambda, \alpha, \beta, \delta, \mu) = \int \mathrm{gh}(x - y)\,\mathrm{gh}(y)\,dy \\
&= a^2 \int (\delta^2 + (x - y + \mu)^2)^{(\lambda - 1/2)/2} \\
&\quad \times \mathrm{K}_{\lambda - 1/2}\left(\alpha\sqrt{\delta^2 + (x - y + \mu)^2}\right) e^{\beta(x - y - \mu)} \\
&\quad \times (\delta^2 + (y - \mu)^2)^{(\lambda - \frac{1}{2})/2}\,\mathrm{K}_{\lambda - \frac{1}{2}}\left(\alpha\sqrt{\delta^2 + (y - \mu)^2}\right) e^{\beta(y - \mu)}\,dy \\
&= a^2(\lambda, \alpha, \beta, \delta) \exp(\beta x - 2\beta\mu) a^{-2}(\lambda, \alpha, 0, \delta)\,\mathrm{gh}^{*2}(x; \lambda, \alpha, 0, \delta, \mu) \\
&\quad \times \frac{a^2(\lambda, \alpha, \beta, \delta)}{a^2(\lambda, \alpha, 0, \delta) \exp(2\beta\mu)} exp(\beta x)\,\mathrm{gh}^{*2}(x; \lambda, \alpha, 0, \delta, \mu) \\
&= \big(\underbrace{M(\beta; \lambda, \alpha, 0, \delta, \mu)}_{=:M_0(\beta)}\big)^{-2} \exp(\beta x)\,\mathrm{gh}^{*2}(x; \lambda, \alpha, 0, \delta, \mu)
\end{aligned}
$$

Clearly, this result translates to $\mathrm{I\!N}$.

$$
\mathrm{gh}^{*n}(x) = M_0^{-n}(\beta) \exp(\beta x)\,\mathrm{gh}^{*n}(x; \lambda, \alpha, 0, \delta, \mu) \quad \forall n \in \mathrm{I\!N}.
$$

We use the following abbreviation for an element of the convolution semigroup of symmetric GH distributions.

$$
\widetilde{f}(x) = \frac{\exp(\beta x)}{M_0^t(\beta)}\,\mathrm{gh}^{*t}(x; \lambda, \alpha, 0, \delta, \mu), \quad \text{for } 0 < t < 1.
$$

Then we obtain for $0 < s, t < 1$

$$
\begin{aligned}
(\widetilde{f}_s * \widetilde{f}_t)(x) &= \int \frac{e^{\beta(x - y)}}{M_0^t(\beta)}\,\mathrm{gh}^{*t}(x - y; \lambda, \alpha, 0, \delta, \mu)\frac{e^{\beta y}}{M_0^s(\beta)}\,\mathrm{gh}^{*s}(y; \lambda, \alpha, 0, \delta, \mu)\,dy \\
&= \frac{e^{\beta x}}{M_0^{t+s}} \int \mathrm{gh}^{*t}(x - y; \lambda, \alpha, 0, \delta, \mu)\,\mathrm{gh}^{*s}(y; \lambda, \alpha, 0, \delta, \mu)\,dy \\
&= \frac{e^{\beta x}}{M_0^{t+s}(\beta)}(\mathrm{gh}^{*t}(\cdot; \lambda, \alpha, 0, \delta, \mu) * \mathrm{gh}^{*s}(\cdot; \lambda, \alpha, 0, \delta, \mu))(x) \\
&= \frac{e^{\beta x}}{M_0^1(\beta)}\,\mathrm{gh}(x; \lambda, \alpha, 0, \delta, \mu).
\end{aligned}
$$

At last we define $\mathrm{gh}^{*t}(x) = \widetilde{f}(x)$, which yields the result. □

Consequently, we obtain a real valued expression for the characteristic function in the symmetric centered case $\beta = \mu = 0$.

$$
\varphi(u; \lambda, \alpha, \delta) = \left(\frac{\alpha^2}{\alpha^2 + u^2}\right)^{\lambda/2} \frac{\mathrm{K}_\lambda(\delta\sqrt{\alpha^2 + u^2})}{\mathrm{K}_\lambda(\alpha\delta)}, \tag{1.45}
$$

and the density $\mathrm{gh}^{*t}$ of $\mathcal{L}(X_t)$ with $X_t$ symmetric and centered, is given by the Fourier inversion formula

$$
\mathrm{gh}^{*t}(x; \lambda, \alpha, 0, \delta, 0) = \frac{1}{\pi} \int_0^\infty \cos(ux)\varphi(u; \lambda, \alpha, \delta)^t\,du \tag{1.46}
$$

25

where $t > 0$. The Fast Fourier Transformation (FFT) allows to compute this density in an efficient way (Press, Teukolsky, Vetterling, and Flannery 1992). At the end of 1.14 we propose an alternative method to compute quantiles of convolutions of GH distributions, which translates also to densities. Of course, in the case of the NIG distribution the density gh$^{*t}$ can be computed directly

$$\text{nig}^{*t}(x; \alpha, \beta, \delta, 0) = \text{nig}(x; \alpha, \beta, t\delta, 0) \quad \text{for } t > 0. \tag{1.47}$$

On the one hand the estimation of NIG distributions is more time consuming, on the other hand convolutions, necessary for option pricing, are easier to obtain.

## 1.11 The Exponential Lévy Model

We consider the following financial market living on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ which satisfies the usual conditions. We model two assets: The first one is a risk free asset $(B_t)_{0 \leq t \leq T}$ evolving according to the equation

$$\begin{cases} \mathrm{d}B_t = rB_t\mathrm{d}t, & 0 \leq t \leq T \\ B_0 = 1, \end{cases} \tag{1.48}$$

where $r$ is a given interest rate and $T$ is the fixed horizon date for all market activities. The second asset is a stock with price process $(S_t)_{0 \leq t \leq T}$ given as follows

$$S_t = S_0 \exp(X_t). \tag{1.49}$$

where $X_t$ is a GH Lévy process. This process is purely discontinuous as well. We are in an incomplete setting, therefore we have to select a particular equivalent martingale measure. Arbitrage free prices are obtained as expectations under these measures (Delbaen and Schachermayer 1994). Moreover, Eberlein and Jacod (1997) have shown in the hyperbolic model that prices in the whole trivial no-arbitrage-interval are possible. See Raible (1999) for similar results for other GH distributions.

## 1.12 Derivative Pricing by Esscher Transforms

In the sequel we investigate the risk-neutral probability measure obtained as Esscher equivalent measure $P^\vartheta$ given by

$$dP^\vartheta = e^{\vartheta X_t - t \log M(\vartheta)} dP. \tag{1.50}$$

Sometimes (1.50) is called exponential tilting. Esscher (1932) introduced these measure transforms for different economic and statistical applications (see for instance Jensen 1995, pp. 11–16 for a discussion). Gerber and Shiu (1994) proposed to use this equivalent measure transformation for the pricing of derivatives. They define the Esscher transform for $0 \leq t \leq \infty$. This transformation does not define an equivalent measure on $\mathcal{F}_\infty^X$ (see Proposition 19 Remark ii in Keller 1997). Otherwise, if we look only on a finite time horizon the problem is not to proof the equivalence of the measure but to give a consistency condition, that the new process is an equivalent Lévy process (see the remark of Delbaen, Schachermayer and Schweizer in the discussion of Gerber and Shiu 1994, pp. 148–151).

It is remarkable that the Esscher tranform may also be obtained by a transformation of the Lévy triplet.

**Lemma 1.51.** *Let $X$ be a Lévy process with Lévy triplet $(0, 0, F)$ and let $y(x) = \exp(\vartheta x)$ and suppose that $\int_{|x| \geq 1} \exp(\vartheta x) dF < \infty$, then the new measure $Q$ defined by the change of triplet according to*

$$\left( b' = b + \int h(y-1)df, c' = 0, F' = yF \right)$$

*is the Esscher transform of $P$.*

The assertion was proved by Keller (1997, Lemma 21 and equation 1.29). The lemma allows to compute the Esscher transformed risk-neutral density by Fourier inversion methods when only the characteristic function of the real-world measure is known.

## Esscher Transforms in Finance

In a discrete time setting a generalization of (1.50) is proposed by Bühlmann, Delbaen, Embrechts, and Shiryaev (1996) considering conditional Esscher transforms. In their case one finds a reason to choose the Esscher transform as a risk-neutral measure by equilibrium arguments. See also Shiryaev (1999, pp. 683–690) for Esscher transforms in finance and Keller (1997, Section 1.4.3) for an equilibrium approach leading to Esscher transform as a risk-neutral measure. However, in the two equilibrium approaches the settings and utility functions are different.

Moreover, Grandits (1996) considered one step processes: then the risk-neutral Esscher probability measure corresponds to the $p$-optimal measure for $p \to 1$. The $p$-optimal measure is a generalization of the variance-optimal measure.

From a mathematical and numerical point of view the simplicity and analytical tractability of the Esscher transform is favourable.

## Esscher Prices and the Minimal Martingale Measure

Chan (1999) asserts that the Esscher transform yields an equivalent martingale measure for the stockprice model

$$\mathrm{d}S_t = \sigma_t S_{t-} \mathrm{d}X_t + b_t S_{t-} \mathrm{d}t \tag{1.52}$$

where $(X_t)_{0 \leq t \leq T}$ is a Lévy process with $E[\exp(hX_1)] < \infty$ for all $h \in (h_1, h_2)$, $0 < h_1, h_2 \leq \infty$. The diffusion and drift coefficients $(\sigma_t)_{0 \leq t \leq T}$ and $(b_t)_{0 \leq t \leq T}$ are defined to be deterministic continuous functions. The condition on $X$ in Chan (1999), which is fulfilled for a GH Lévy motion, implies that $X_t$ has finite moments of all order. In the model (1.52) the Esscher transform yields a minimal martingale measure in the sense of Föllmer and Schweizer (1991).

On the contrary, in the model

$$S_t = S_0 \exp(\sigma_t X_t + b_t t), \tag{1.53}$$

where the price follows an exponential Lévy process, the minimal martingale measure does not coincide with the Esscher transform. However, both measures are very close. They are identical in the case of a Brownian motion $X$, and if the Lévy measure of $X$ is concentrated around 0, the Esscher martingale measure approximates the minimal martingale measure.

## Esscher Prices in the GH Model

With Keller (1997, Lemma 14) we have to find the solution of

$$r = \log \frac{M(\vartheta + 1)}{M(\vartheta)} \tag{1.54}$$

$$= \mu + \frac{\lambda}{2} \log \frac{\alpha^2 - (\beta + \vartheta)^2}{\alpha^2 - (\beta + \vartheta + 1)^2} + \log \frac{K_\lambda \left( \delta \sqrt{\alpha^2 - (\beta + \vartheta)^2} \right)}{K_\lambda \left( \delta \sqrt{\alpha^2 - (\beta + \vartheta + 1)^2} \right)},$$

where $M$ is the moment generating function given in Lemma 1.13 and $r > 0$ the constant interest rate (corresponding to the time lag for which the GH law was estimated). We find the solution $\vartheta_0$ by a refined bracketing method. In the case of a NIG distribution equation (1.54) simplifies to

$$r = \mu + \delta\left( \sqrt{\alpha^2 - (\beta + \vartheta)^2} - \sqrt{\alpha^2 - (\beta + \vartheta + 1)^2} \right). \tag{1.55}$$

The density of the Esscher transform is obtained from the convolution semi-group $(\mathrm{gh}^{*t})_{t>0}$. To avoid the computation of convolutions for asymetric densities we reduce the computation, in the way we described above, with FFT to the symmetric case. We use the fact that Lemma 52 of Keller (1997) translates directly to generalized hyperbolic distributions. $M_0(u)$ denotes the moment generating function given in Definition 1.13 for $\beta = 0$.

$$\mathrm{gh}^{*t,\vartheta}(x; \lambda, \alpha, \beta, \delta, \mu) = \frac{e^{\vartheta x}}{M^t(\vartheta)} \, \mathrm{gh}^{*t}(x; \lambda, \alpha, \beta, \delta, \mu)$$

$$= \frac{e^{\vartheta x}}{M^t(\vartheta)} \frac{e^{\beta x}}{M_0^t(\beta)} \, \mathrm{gh}^{*t}(x - \mu t; \lambda, \alpha, 0, \delta, 0). \tag{1.56}$$

In the case of the NIG distribution we obtain

$$\mathrm{gh}^{*t,\vartheta}(x; -1/2, \alpha, \beta, \delta, \mu) = \frac{e^{\vartheta x}}{M^t(\vartheta)} \, \mathrm{nig}(x; \alpha, \beta, t\delta, t\mu) \tag{1.57}$$

and especially in the case of a centered NIG distribution we obtain a risk-neutral NIG distribution

$$\mathrm{gh}^{*t,\vartheta}(x; -1/2, \alpha, \beta, \delta, 0) = \mathrm{nig}(x; \alpha, \beta + \vartheta, t\delta, 0). \tag{1.58}$$

Consequently, centered NIG transforms are the easiest GH Esscher transforms to implement.

The price of a derivative contract with payoff function $h(S_T)$ is (for all GH distributions) given by

$$E_{Q^\vartheta}[e^{-rT} h(S_T)] = e^{-rT} \int h(S_0 e^x) dQ^\vartheta_{X_T},$$

where $Q^\vartheta$ and $Q^\vartheta_{X_T}$ denote risk-neutral Esscher transformed probability measure and distribution of $X_T$ respectively. For the price of an European call option with payoff function $h(S_T) = (S_T - K)^+$ we obtain

$$C_{\mathrm{gh}} = E_{Q^{\vartheta_0}}[e^{-rT} h(S_T)] = e^{-rT} E_{Q^{\vartheta_0}}[(S_T - K)^+] \tag{1.59}$$

$$= e^{-rT} E_{Q^{\vartheta_0}}[(S_0 e^{X_T} - K)^+] = e^{-rT} \int_{\{S_0 e^x > K\}} (S_0 e^x - K) dQ^{\vartheta_0}_{X_T}(x)$$

$$= S_0 \int_{\log \frac{K}{S_0}}^\infty \mathrm{gh}^{*T,\vartheta_0+1}(x; \dots) dx - e^{-rT} K \int_{\log \frac{K}{S_0}}^\infty \mathrm{gh}^{*T,\vartheta_0}(x; \dots) dx.$$

For the last step we follow Keller (1997, Section 1.4.2).

## 1.13 Computation of Esscher Prices by FFT

The numerical computation of option prices by fast Fourier transforms (FFT) procedes in four steps:

1. First, we solve equation (1.54), or the given simpler versions for the various subclasses, to obtain $\vartheta_0$ by a refined bracketing method. For $\vartheta_0$ we have the following condition $-\alpha - \beta < \vartheta_0 < \alpha - \beta - 1$. We use these bounds as starting values for the Illinois method (Thisted 1988). This approach allows us to compute $\vartheta_0$ in less than 10 iterations with an error smaller than $10^{-10}$.

2. Second, we calculate the symmetric density $\mathrm{gh}^{*T}(x - \mu t; \alpha, 0, \delta, 0)$ using the Fourier inversion formula (1.46). Although the implementation of sophisticated FFT algorithms needs some attention to get exact results, the execution needs only fractions of a second.

   Two-fold convolutions are used as a standard method in the field of image processesing and filtering. The numerical problems of the discrete Fourier transforms and simulation results are treated extensively by the computer scientists. Although characteristic functions are often used in probability theory to prove limit theorems, numerical algorithms for $n$-fold convolutions with $n > 2$ for arbitrary measures are no subject in the numerical literature.

   Unfortunately, only in the case of the NIG subclass the density is available as a closed formula.

3. The Esscher transform is calculated as given in equation (1.56). Since $\beta$ and $\mu$ have often opposite sign, it is numerically favourable to "skew" the distribution in one step.

4. The last step is the calculation of the prices with formula (1.59), which is essentially a numerical integration.

We summarize the numerical results for the comparison of GH Esscher prices with Black-Scholes prices in Section 2.3.

## 1.14 Saddle-Point Approximation to Esscher Prices

An alternative approach to calculate Esscher prices is based on the saddle-point method, a well-known tool in statistics to approximate densities or tail probabilities, if only the moment-generating or cumulant generating function is known. See Jensen (1995) for a thorough exposition of saddle-point approximation methods. The idea behind saddle-point approximations is to compute the density or a tail probability by the Fourier inversion formula (see equation 1.46 for a special case), not by discretizing and the application of the FFT algorithm, but by an expansion of the convex cumulant generating function at the saddle-point. Hence, the first derivative vanishes and the approximation obtains a simple form. Rogers and Zane (1999) proposed to compute option prices by saddle-point methods. In particular, they propose this for stock price models in which $X_T$ has a NIG distribution[6] under the martingale measure—avoiding any "soul-searching" to find a particular martingale measure. In the sequel we show

---

[6]Rogers and Zane (1999) confuse NIG and hyperbolic distributions. Actually, they give the characteristic function of the NIG distribution. Since convolutions of NIG distributions are known explicitly, it is not necessary to approximate them by saddle-point methods.

that this approximation method is also useful for the calculation of prices with the Esscher change of measure.

As a first step, we compute the cumulant generating function of $X_T$ under the Esscher equivalent martingale measure where $(X_t)_{0 \le t \le T}$ follows a GH Lévy process. The *cumulant generating function* (CGF) is defined as the log of the moment generating function (MGF)

$$\kappa(u) := \log M(u) = \log E[\exp(uX)]. \tag{1.60}$$

Let $M(u,t)$ and $\kappa(u,t)$ denote the MGF and CGF of the $t$-fold convolution. Since $M(u,t) = M(u)^t$ for the convolution of i.i.d. random variates, we also have $\kappa(u,t) = t\kappa(u)$.

The moment generating function of the Esscher transform with parameter $\vartheta$ is given by (Gerber and Shiu 1994)

$$M(u,t,\vartheta) = \frac{M(u+\vartheta,t)}{M(u,t)}, \tag{1.61}$$

where $\vartheta \in \mathbb{R}$ is the parameter of the Esscher transform. Using Lemma 1.13 we obtain for the CGF of the Esscher transform in the GH model:

$$
\begin{aligned}
\kappa(u,t,\vartheta) &:= \log M(u,t,\vartheta) = \log \frac{M(u+\vartheta,t)}{M(u,t)} \\
&= \log \frac{M(u+\vartheta)^t}{M(u)^t} = t \log \frac{M(u+\vartheta)}{M(u)} \\
&= t \log \left[ \frac{e^{(u+\vartheta)\mu}}{e^{\vartheta\mu}} \left( \frac{\alpha^2 - (\beta+\vartheta)^2}{\alpha^2 - (\beta+u+\vartheta)^2} \right)^{\lambda/2} \frac{K_\lambda \left( \delta\sqrt{\alpha^2 - (\beta+u+\vartheta)^2} \right)}{K_\lambda \left( \delta\sqrt{\alpha^2 - (\beta+\vartheta)^2} \right)} \right] \\
&= tu\mu + \frac{t\lambda}{2} \log \frac{\alpha^2 - (\beta+\vartheta)^2}{\alpha^2 - (\beta+u+\vartheta)^2} + \log \frac{K_\lambda \left( \delta\sqrt{\alpha^2 - (\beta+u+\vartheta)^2} \right)}{K_\lambda \left( \delta\sqrt{\alpha^2 - (\beta+\vartheta)^2} \right)}.
\end{aligned}
$$

Note, that $\kappa(u,t\vartheta)$ exists for $|\beta + u + \vartheta| < \alpha$. Since $\kappa(u,t,\vartheta)$ is the cumulant generating function of a $t$-fold convolution of a GH distribution with parameter $\beta' = \beta + \vartheta$, we have shown again that the risk-neutral Esscher transformation of a Lévy process changes only the skewness in the following way $\beta \mapsto \beta + \vartheta^*$.

The price of a European *put option* under an equivalent martingale measure $Q$ is according to the martingale pricing theory given by

$$p = E^Q[e^{-rT}(K - S_T)^+] \tag{1.62}$$

with strike price $K$ and stock price $S_T$ under $Q$. Since we model the logarithm $X_t = \log(S_t/S_0)$ of the stockprice $S_t$, we could also write

$$
\begin{aligned}
p &= e^{-rT} E^Q[(K - S_0 e^{X_T})^+] \\
&= e^{-rT + \log S_0} E^Q \left[ \left( \exp(\underbrace{\log K - \log S_0}_{=:\, k}) - \exp(X_T) \right)^+ \right] \\
&= S_0 e^{-rT} E^Q \left[ \left( e^k - e^{X_T} \right)^+ \right], \tag{1.63}
\end{aligned}
$$

where $k = \log(K/S_0)$. Note, that the value of a European put $p$ and a European call $c$ with the same strike price $K$ are closely related by

$$p = c - S_0 + e^{-rT}K, \tag{1.64}$$

where $S_0$ is the stock price at time 0. This follows from a trivial arbitrage consideration and does not depend on the model.

Now we construct an exponential family by a second exponential tilting under the martingale measure $Q$. See Jensen (1995, Chapter 1.2) for a short introduction.

We assume that the cumulant generating function $\kappa$ of $X_T$ under $Q$ is finite in some open stripe $\{z : -a_- < \mathrm{Re}\, z < a_+\}$ containing the imaginary axis with $a_- > 0$ and $a_+ > 1$. Define the measure $Q_y$ by the

$$\frac{dQ_y}{dQ} = \exp\big(yX_T - \kappa(y)\big). \tag{1.65}$$

Note, that this exponential tilting (or Esscher transform) is not related to the Esscher tranform used to obtain the equivalent martingale measure. Then the cumulant transform $\kappa_y$ of $Q_y$ is given by

$$\kappa_y(u) = \kappa(y + u) - \kappa(y).$$

Now we can rewrite (1.63) as

$$p = S_0 e^{k - rT} Q(X_T < k) - S_0 e^{-rT + \kappa(1)} Q_1(X_T < k). \tag{1.66}$$

The tail probabilities $Q(X_T < k)$ and $Q_1(X_T < k)$ may now be approximated by saddle-point methods.

## Lugannani-Rice Formula

We start by reviewing one particular saddle-point method. The exponential family generated by $X$ and $Q$ defined in (1.65) has the following mean and variance

$$\mu(y) = E_y X = \frac{\mathrm{d}\kappa}{\mathrm{d}y}(y) \tag{1.67}$$

$$\sigma(y) = \mathrm{Var}_y X = \frac{\mathrm{d}^2 \kappa}{\mathrm{d}y\mathrm{d}y}(y). \tag{1.68}$$

The solution $\hat{y}$ of $E_y(X) = x$, $x \in \mathrm{int}\, C$, where $C$ is the closed convex hull of the support of $X$, is unique. When the exponential family is steep we have in fact that $\{E_y X : y \in \mathrm{int}\, C\} \subset C$, i.e. $\mu(y) = x$ has a solution $\hat{y} \in \mathrm{int}\, \Theta := \{y \in \mathbb{R} : \exp(\kappa(y)) < \infty\}$ for any $x \in \mathrm{int}\, C$ (Barndorff-Nielsen 1978, Section 9.3).

One application of saddle-point methods is to find approximations to sums $\overline{X}_n$ of $X_1, \ldots, X_n$ i.i.d. random variates. The following formula for the approximation of the lower tail probability proved by Lugannani and Rice (1980) is an improvement to other saddle-point approximations, since it is valid for all $x \in \mathbb{R}$. See Daniels (1987) for a comparison of the Lugannani-Rice approach to other saddle-point methods.

**Theorem 1.69 (Lugannani-Rice Formula).** *Assume that the solution $\hat{y}$ exists and Laplace-transform integrable. Then*

$$P(\overline{X}_n \le x) = \Phi\big(r(x)\big) - \varphi\big(r(x)\big)\left(\frac{1}{\lambda(x)} - \frac{1}{r(x)} + O(n^{-3/2})\right),$$

*where $\varphi$ and $\Phi$ are the density and CDF function of the standard normal distribution, respectively. Furthermore*

$$r(x) = \text{sign}\big(\hat{y}(x)\big)\sqrt{2n\big(\hat{y}(x)x - \kappa(\hat{y}(x))\big)} \qquad (1.70)$$

$$\lambda(x) = \sqrt{n}\,\hat{y}(x)\,\sigma(\hat{y}(x)). \qquad (1.71)$$

See Jensen (1995, 3.3.2-9) for the proof.

It is remarkable that under the given assumptions the Lugannani-Rice formula holds uniformly for $\hat{y}(x)$ in a compact subset of int $\Theta$ (Jensen 1995, Theorem 3.3.2). Note that a better approximation may be achieved using the higher order terms given in the original paper of Lugannani and Rice (1980). In general, simulation studies indicate that the use of the first order approximation yields good results (Daniels 1987).

### Properties of the Cumulant Generating Function

Finally, we have to present the properties of GH distributions necessary for the application of Theorem 1.69.

- The CGF of the GH distribution is the logarithm of the moment generating function given in Lemma 1.13. For the parameter values which were typically estimated from financial time series, the MGF and hence the CGF exists and fulfills $a_+ > 1$.

- The Laplace transform is integrable (this follows from Lemma 1.13 and Lukasz 1970, Section 7.2).

- With Lemma 1.13, equation (1.17) and Definition B.26 the we can compute the first derivative of the CGF $\kappa(u)$. We assume $\mu = 0$ for simplicity and obtain

$$\kappa'(y) = \frac{R_\lambda\big(\delta f(u)\big)}{f(y)} \qquad \text{for } |\beta + y| < \alpha, \qquad (1.72)$$

where $f(y) = \sqrt{\alpha^2 - (\beta + y)^2}$. The second derivative is easily computed with equation (B.32).

Hence, if a solution $\hat{y}(x)$ exists, we can apply the Lugannani-Rice Formula to compute the option price given in equation (1.66).

Note, that the approximation by saddle-point method, which we have proposed here, is also valid for the approximation of quantiles of convolutions of Gh distributions. Consider, for example, the calculation of the VaR for a holding period of ten days from a hyperbolic distribution estimated from daily log-returns.

## 1.15 Conclusion

In this chapter we developed an algorithm to estimate parameters for the class of generalized hyperbolic distributions which includes the hyperbolic and the normal inverse Gaussian distribution as special cases. We compared the results of the estimations for financial return data sets. In general, generalized hyperbolic distributions and their subclasses provide better fits to the data than the normal distribution. As expected, the best fits are obtained for the generalized hyperbolic distributions followed by the NIG and the hyperbolic distributions.

Table 1.11: *Computation times: GH and hyperbolic prices are computed by FFT algorithms.*

|  | 1000 call prices | | 1000 estimations | |
|---|---|---|---|---|
| Black/Scholes |  | 3 sec |  |  |
| GH | 2 min | 28 sec | 2 h | 16 min |
| NIG |  | 6 sec | 2 h | 15 min |
| hyperbolic |  | 15 sec | 1 h | 31 min |

It is worth to mention that generalized hyperbolic distributions tend to overfitting and that the estimation is computationally demanding. The computation times necessary for estimation and derivative pricing are given in Table 1.11. Hyperbolic and NIG distributions provide an acceptable tradeoff between the accuracy of fit and and the necessary numerical effort.

## 1.16 Tables

Table 1.12: *Minimum distance estimates for Deutsche Bank returns.*

| $\lambda$ | $\alpha$ | $\beta$ | $\delta$ | $\mu$ | Distance | $\chi$ | $\xi$ | $\varrho$ | $\zeta$ |
|---|---|---|---|---|---|---|---|---|---|
| *Maximum Likelihood* | | | | | | | | | |
| -1.0024 | 39.6 | 4.14 | 0.0118 | -0.000158 | 4878.00 | 0.086 | 0.827 | 0.104 | 0.463 |
| NIG | 59.4 | 4.64 | 0.0094 | -0.000226 | 4877.62 | 0.063 | 0.802 | 0.078 | 0.556 |
| HYP | 114.8 | 3.35 | 0.0000 | -0.000000 | 4872.25 | 0.029 | 1.000 | 0.029 | 0.000 |
| *Minimal Kolmogorov Distance* | | | | | | | | | |
| -0.5002 | 63.8 | 3.81 | 0.0097 | -0.000211 | 0.013469 | 0.047 | 0.786 | 0.060 | 0.620 |
| NIG | 63.8 | 3.81 | 0.0097 | -0.000211 | 0.013470 | 0.047 | 0.786 | 0.060 | 0.621 |
| HYP | 116.8 | 4.60 | 0.0000 | -0.000211 | 0.014424 | 0.039 | 1.000 | 0.039 | 0.000 |
| *Minimal Anderson & Darling Statistic* | | | | | | | | | |
| -0.7162 | 39.6 | 4.00 | 0.0117 | -0.000158 | 0.10 | 0.083 | 0.827 | 0.101 | 0.463 |
| NIG | 48.5 | 4.05 | 0.0118 | -0.000158 | 0.13 | 0.067 | 0.799 | 0.084 | 0.568 |
| HYP | 80.5 | 2.98 | 0.0000 | -0.000162 | 0.20 | 0.037 | 1.000 | 0.037 | 0.000 |
| *Minimal $L^1$ Distance* | | | | | | | | | |
| 0.0590 | 85.5 | 5.62 | 0.0073 | -0.000282 | 0.000352 | 0.052 | 0.786 | 0.066 | 0.620 |
| NIG | 64.9 | 3.79 | 0.0098 | -0.000064 | 0.000337 | 0.046 | 0.782 | 0.058 | 0.636 |
| HYP | 116.5 | 6.12 | 0.0000 | -0.000328 | 0.000407 | 0.053 | 1.000 | 0.053 | 0.000 |
| *Minimal $L^2$ Distance* | | | | | | | | | |
| 0.4900 | 102.7 | 7.24 | 0.0052 | -0.000459 | 0.00111 | 0.057 | 0.807 | 0.070 | 0.536 |
| NIG | 64.2 | 6.48 | 0.0098 | -0.000382 | 0.00119 | 0.079 | 0.785 | 0.101 | 0.623 |
| HYP | 122.7 | 7.68 | 0.0022 | -0.000503 | 0.00114 | 0.056 | 0.887 | 0.063 | 0.270 |

Table 1.13: *Maximum likelihood estimates of generalized hyperbolic distributions and the following subclasses: λ= –3/2 (in Equation 1.23), NIG (λ= –1/2), hyperboloid (λ=0), hyperbola (λ=1/2), and hyperbolic (λ=1) distributions. New York Stock Exchange Indices from January 2, 1990 to November 29, 1996. The first line always contains the estimate for arbitrary λ. 1746 observations.*

| $\lambda$ | $\alpha$ | $\beta$ | $\delta$ | $\mu$ | LogLH | $\chi$ | $\xi$ | $\varrho$ | $\zeta$ |
|---|---|---|---|---|---|---|---|---|---|
| *NYSE Composite Index* | | | | | | | | | |
| 0.83 | 213.74 | −6.22 | 0.0022 | 0.00066 | 6396.06 | -0.024 | 0.83 | -0.029 | 0.47 |
| −1.5 | 74.9 | −10.44 | 0.0084 | 0.00085 | 6392.09 | -0.11 | 0.78 | -0.14 | 0.62 |
| −0.5 | 136.29 | −8.95 | 0.0059 | 0.00079 | 6394.56 | -0.049 | 0.74 | -0.066 | 0.81 |
| 0 | 164.91 | −8.06 | 0.0046 | 0.00075 | 6395.39 | -0.037 | 0.75 | -0.049 | 0.76 |
| 0.5 | 193.86 | −6.99 | 0.0032 | 0.0007 | 6395.93 | -0.028 | 0.78 | -0.036 | 0.62 |
| 1 | 225.03 | −5.84 | 0.0016 | 0.00065 | 6396.01 | -0.022 | 0.86 | -0.026 | 0.35 |
| *NYSE FINANCE Index* | | | | | | | | | |
| 0.05 | 151.55 | −4.57 | 0.0062 | 0.00074 | 6067.61 | -0.022 | 0.72 | -0.03 | 0.94 |
| −1.5 | 73.07 | −3.96 | 0.01 | 0.00071 | 6066.60 | -0.041 | 0.75 | -0.054 | 0.77 |
| −0.5 | 125.26 | −4.28 | 0.0078 | 0.00073 | 6067.49 | -0.024 | 0.71 | -0.034 | 0.97 |
| 0 | 149.09 | −4.54 | 0.0063 | 0.00074 | 6067.61 | -0.022 | 0.72 | -0.03 | 0.95 |
| 0.5 | 172.48 | −4.91 | 0.0048 | 0.00076 | 6067.53 | -0.021 | 0.74 | -0.028 | 0.83 |
| 1 | 196.24 | −5.46 | 0.003 | 0.0008 | 6067.23 | -0.022 | 0.79 | -0.028 | 0.59 |
| *NYSE Industrial Index* | | | | | | | | | |
| 1.36 | 243.92 | −10.37 | 8.9e-10 | 0.0009 | 6317.80 | -0.043 | 1 | -0.043 | 2.2e-07 |
| −1.5 | 75.42 | −10.26 | 0.0089 | 0.00091 | 6317.36 | -0.11 | 0.78 | -0.14 | 0.66 |
| −0.5 | 135.3 | −9.39 | 0.0064 | 0.00087 | 6319.21 | -0.051 | 0.73 | -0.069 | 0.86 |
| 0 | 163.5 | −9.02 | 0.0051 | 0.00085 | 6319.61 | -0.041 | 0.74 | -0.055 | 0.83 |
| 0.5 | 192.02 | −8.76 | 0.0037 | 0.00083 | 6319.61 | -0.035 | 0.76 | -0.046 | 0.72 |
| 1 | 222.25 | −8.83 | 0.0022 | 0.00083 | 6319.05 | -0.032 | 0.82 | -0.04 | 0.5 |
| *NYSE Transport Index* | | | | | | | | | |
| −2.31 | 7.19 | 7.19 | 0.016 | -0.00028 | 5725.53 | 1 | 1 | 1 | 2e-06 |
| −1.5 | 62.99 | 7.4 | 0.013 | -0.0003 | 5724.14 | 0.087 | 0.74 | 0.12 | 0.81 |
| −0.5 | 109.49 | 7.74 | 0.01 | -0.00032 | 5721.97 | 0.049 | 0.69 | 0.071 | 1.1 |
| 0 | 129.88 | 7.87 | 0.0084 | -0.00033 | 5720.77 | 0.042 | 0.69 | 0.061 | 1.1 |
| 0.5 | 149.56 | 8.01 | 0.0068 | -0.00034 | 5719.46 | 0.038 | 0.71 | 0.054 | 1 |
| 1 | 169.14 | 8.2 | 0.005 | -0.00035 | 5717.98 | 0.036 | 0.74 | 0.048 | 0.84 |
| *NYSE Utility Index* | | | | | | | | | |
| 1.84 | 298.21 | −9.98 | 3e-11 | 0.00055 | 6379.26 | -0.033 | 1 | -0.033 | 8.9e-09 |
| −1.5 | 122.16 | −9.16 | 0.0095 | 0.00052 | 6380.93 | -0.051 | 0.68 | -0.075 | 1.2 |
| −0.5 | 180.05 | −9.17 | 0.0075 | 0.00052 | 6380.88 | -0.033 | 0.65 | -0.051 | 1.4 |
| 0 | 206.46 | −9.22 | 0.0065 | 0.00052 | 6380.72 | -0.029 | 0.65 | -0.045 | 1.3 |
| 0.5 | 232 | −9.31 | 0.0053 | 0.00052 | 6380.48 | -0.027 | 0.67 | -0.04 | 1.2 |
| 1 | 257.19 | −9.44 | 0.0041 | 0.00053 | 6380.15 | -0.026 | 0.7 | -0.037 | 1 |

Table 1.14: *Maximum likelihood estimates of generalized hyperbolic distributions and the following subclasses: λ= –3/2 (in Equation 1.23), NIG (λ= –1/2), hyperboloid (λ=0), hyperbola (λ=1/2), and hyperbolic (λ=1) distributions. German stocks from January 1988 to May 1994. The first line always gives the estimate for arbitrary λ. 1598 observations.*

| λ | α | β | δ | μ | LogLH | χ | ξ | ϱ | ζ |
|---|---|---|---|---|---|---|---|---|---|
| *Bayer* | | | | | | | | | |
| −1.79 | 21.3 | 2.67 | 0.0153 | -0.000004 | 5003.69 | 0.109 | 0.869 | 0.125 | 0.323 |
| −1.5 | 37.58 | 2.87 | 0.014 | -2.7e-05 | 5003.52 | 0.062 | 0.81 | 0.076 | 0.53 |
| NIG | 81.6 | 3.69 | 0.0103 | -0.000123 | 5001.54 | 0.033 | 0.737 | 0.045 | 0.843 |
| 0 | 101.03 | 4.17 | 0.0085 | -0.00018 | 5000.05 | 0.03 | 0.73 | 0.041 | 0.86 |
| 0.5 | 119.93 | 4.72 | 0.0065 | -0.00024 | 4998.23 | 0.029 | 0.75 | 0.039 | 0.78 |
| HYP | 139.0 | 5.35 | 0.0044 | -0.000311 | 4996.00 | 0.030 | 0.789 | 0.039 | 0.608 |
| *Bayerische Hypotheken und Wechselbank* | | | | | | | | | |
| −1.59 | 17.9 | 2.19 | 0.0157 | -0.000072 | 4815.05 | 0.108 | 0.884 | 0.122 | 0.278 |
| −1.5 | 22.82 | 2.29 | 0.015 | -8.6e-05 | 4815.03 | 0.087 | 0.86 | 0.1 | 0.35 |
| NIG | 63.8 | 3.12 | 0.0106 | -0.000211 | 4813.37 | 0.038 | 0.773 | 0.049 | 0.674 |
| 0 | 81.68 | 3.48 | 0.0083 | -0.00026 | 4811.72 | 0.033 | 0.77 | 0.043 | 0.68 |
| 0.5 | 99.58 | 3.8 | 0.006 | -0.00031 | 4809.41 | 0.03 | 0.79 | 0.038 | 0.6 |
| HYP | 118.5 | 4.03 | 0.0035 | -0.000330 | 4806.13 | 0.029 | 0.840 | 0.034 | 0.418 |
| *Daimler Benz* | | | | | | | | | |
| −1.68 | 13.0 | 3.93 | 0.0182 | -0.000539 | 4625.48 | 0.272 | 0.903 | 0.301 | 0.227 |
| −1.5 | 21.54 | 4.13 | 0.017 | -0.00058 | 4625.38 | 0.16 | 0.86 | 0.19 | 0.36 |
| NIG | 57.6 | 5.09 | 0.0120 | -0.000748 | 4623.24 | 0.068 | 0.769 | 0.088 | 0.691 |
| 0.5 | 88.88 | 5.99 | 0.0068 | -0.0009 | 4619.3 | 0.053 | 0.79 | 0.067 | 0.6 |
| 0 | 73.31 | 5.52 | 0.0095 | -0.00082 | 4621.52 | 0.058 | 0.77 | 0.075 | 0.69 |
| HYP | 105.2 | 6.58 | 0.0039 | -0.000999 | 4616.28 | 0.053 | 0.843 | 0.063 | 0.406 |
| *Deutsche Bank, see Table 1.12.* | | | | | | | | | |
| *Siemens* | | | | | | | | | |
| −1.89 | 3.680 | 3.65 | 0.0164 | -0.000056 | 4914.74 | 0.988 | 0.996 | 0.992 | 0.008 |
| −1.5 | 28.61 | 3.95 | 0.015 | -9.6e-05 | 4913.72 | 0.12 | 0.84 | 0.14 | 0.41 |
| NIG | 74.7 | 4.76 | 0.0107 | -0.000188 | 4908.68 | 0.047 | 0.745 | 0.064 | 0.800 |
| 0 | 94.27 | 5.06 | 0.0089 | -0.00022 | 4905.66 | 0.04 | 0.74 | 0.054 | 0.84 |
| 0.5 | 113.09 | 5.31 | 0.007 | -0.00025 | 4902.34 | 0.035 | 0.75 | 0.047 | 0.79 |
| HYP | 131.8 | 5.52 | 0.0049 | -0.000266 | 4898.64 | 0.033 | 0.780 | 0.042 | 0.644 |
| *Thyssen* | | | | | | | | | |
| −2.34 | 1.4 | 0.92 | 0.0241 | 0.000545 | 4558.12 | 0.654 | 0.988 | 0.662 | 0.025 |
| −1.5 | 42.92 | 1.27 | 0.02 | 0.00047 | 4557.22 | 0.022 | 0.73 | 0.03 | 0.86 |
| NIG | 72.1 | 1.62 | 0.0154 | 0.000399 | 4555.85 | 0.015 | 0.688 | 0.022 | 1.110 |
| 0 | 84.63 | 1.8 | 0.013 | 0.00036 | 4555.16 | 0.015 | 0.69 | 0.021 | 1.1 |
| 0.5 | 96.4 | 1.96 | 0.01 | 0.00033 | 4554.52 | 0.015 | 0.71 | 0.02 | 0.97 |
| HYP | 107.6 | 2.10 | 0.0066 | 0.000300 | 4553.96 | 0.015 | 0.764 | 0.020 | 0.714 |

Table 1.15: *Comparison of the directly estimated skewness and kurtosis with the skewness and kurtosis calculated from the estimates of GH, NIG and hyperbolic distributions with different metrics (Deutsche Bank and Bayerische Hypotheken und Wechselbank returns).*

| | | Skewness | Kurtosis | Skewness | Kurtosis |
| | | Deutsche Bank | | Bay.Hyp.u.Wechselbank | |
|---|---|---|---|---|---|
| Empirical | | $-0.519$ | 10.872 | $-1.220$ | 15.919 |
| Normal | | 0.0 | 3.0 | 0.0 | 3.0 |
| Maximum Likelihood | GH | 0.378 | 7.492 | 0.291 | 10.413 |
| Maximum Likelihood | NIG | 0.314 | 5.529 | 0.178 | 4.490 |
| Maximum Likelihood | HYP | 0.123 | 3.010 | 0.071 | 3.003 |
| Kolmogorov distance | GH | 0.227 | 4.906 | $-0.793$ | 1.007 |
| Kolmogorov distance | NIG | 0.227 | 4.903 | $-0.000$ | 0.020 |
| Kolmogorov distance | HYP | 0.166 | 3.018 | $-0.009$ | 2.708 |
| Anderson & Darling | GH | 0.419 | 7.233 | 0.058 | 3.210 |
| Anderson & Darling | NIG | 0.332 | 5.427 | $-1.141$ | 9.068 |
| Anderson & Darling | HYP | 0.156 | 3.016 | 0.043 | 2.579 |
| $L^1$ distance | GH | 0.261 | 3.887 | 0.238 | 4.438 |
| $L^1$ distance | NIG | 0.219 | 4.782 | 0.237 | 4.252 |
| $L^1$ distance | HYP | 0.222 | 3.032 | 0.249 | 2.563 |
| $L^2$ distance | GH | 0.281 | 3.315 | 0.323 | 4.087 |
| $L^2$ distance | NIG | 0.383 | 5.010 | 0.320 | 4.024 |
| $L^2$ distance | HYP | 0.246 | 2.764 | 0.227 | 3.034 |

# Chapter 2

# Application and Testing of Derivative Pricing Models

One result of the empirical study in the last chapter is that generalized hyperbolic distributions provide a realistic model for daily stock returns. We have derived an option pricing formula using the Esscher transform as one possibility to find prices in an incomplete market. In this chapter we compare the Black-Scholes model with the GH option pricing model, which is a generalization of the hyperbolic model developed by Eberlein and Keller (1995), Keller (1997).

Moreover, we also propose a general recipe for a comparison of option pricing models with market reality. Some parts of the comparison of the hyperbolic and the Black-Scholes model have already been published in Eberlein, Keller, and Prause (1998).[1] The objectives of our research are to examine

- The consistency of the model assumptions with empirical observed price processes,

- Typical patterns in implicit volatilities[2],

- The application and performance of option pricing models from a practicioner's point of view.

Due to the difficulty if not impossibility to construct a test in a strict statistical sense for option pricing models (see Section 2.7 for some further remarks), we try to understand and compare the models using criteria which are relevant for practice.

On the one hand there is a constant interest in alternative models, on the other hand, from the practicioner's point of view, these models should improve everyday's work. One reason for the widespread use of Black-Scholes type models is that their deficiencies are well-understood. Hence, we compare the advantages and disadvantages of both models, the Black-Scholes and the generalized hyperbolic model. We want to understand the implicit volatility patterns, the pricing behaviour, and the practical value of the generalized hyperbolic alternative to models based on normal distributions. See Rubinstein (1985), Bates (1996), Bakshi, Cao, and Chen (1997) for tests of pricing models driven by Gaussian Lévy processes.

---

[1]Numerical errors are corrected, therefore some results are different.

[2]*Volatility* denotes the annualized standard deviation, obtained from the daily standard deviation under the assumption of independence. We assume 250 trading days per annum.

## 2.1 Preliminaries

### Preparation of Intraday Option Datasets

The study is based on intraday option and stock market data of Bayer, Daimler Benz, Deutsche Bank, Siemens and Thyssen from July 1992 to August 1994. The option data set contains all trades reported by the Deutsche Terminbörse (DTB) during the period above.[3] Per month there are 736 to 7924 observed option and 434 to 2049 observed stock prices.

The data is processed in the following way: At first, we assign to each option price the corresponding intraday stock price. The electronic DTB had longer business hours than the stock exchange in Frankfurt. Following Rubinstein (1985), Clewlow and Xu (1993) we remove all option quotes without stock trading in the preceding 20 minutes. This leads to a removal of approximately half of the option quotes.

The time to maturity is calculated on the basis of actual trading days. This means that the days with trading at the exchanges in Frankfurt during the lifetime of each option were counted. In contrast to Cox and Rubinstein (1985) we use the resulting trading time scale for option pricing and variance estimation. This guarantees comparability between the implicit volatilities produced by contingent claims and the empirical (historical) volatilities of the stock. Trading days are the natural time scale in the hyperbolic case since the model was proposed for the reason of the good fits to daily data.

Dividend payments reduce the price of the stock. Following Kolb (1995) we correct the share value by substracting the discounted dividend, i.e. $S^- = S_0 - d \exp(-rt)$, where $d$ denotes the amount of the dividend payment which is made $t$ trading days after the option trade and $r$ the riskless daily interest rate. On the German market dividends are paid only once a year. Hence, we had to correct about 18% of the values. For the interest rate we used the Frankfurt interbank offered rate (FIBOR) on a monthly basis with different maturities (1, 3, 6 months). Hence, the substantial changes in the term structure in the years from 1992 to 1994 were taken into account. Finally, option prices must satisfy some trivial no-arbitrage conditions (Cox and Rubinstein 1985). If an option quote violates these bounds it is removed. Note, that most of the trading takes place at-the-money and with a short time to maturity ($T = 1, \ldots, 50$ trading days). Therefore, one should pay particular attention to this region.

### Rescaling of Generalized Hyperbolic Distributions

Normal distributions are characterized by the scale and location parameter. The additional parameters of GH distributions allow to specify in particular the tail of the distribution more exactly. Tail estimates in financial applications are typically based on time series observed over a longer horizon. Especially rare events like crashes should be taken into account as accurately as possible. On the contrary, variance estimates should be adatepd regularly to short-term developments. In the case of GH distributions it is possible to distinguish shape parameters from scale and location parameters. We apply this approach frequently: For the calculation of implicit volatilities we rescale the generalized hyperbolic distribution while keeping the shape of the distribution fixed. A similar problem occurs while the calculation

---

[3]Since autumn 1998 Eurex Germany, formerly the DTB (Deutsche Terminbörse), comprises the derivatives market of the Deutsche Börse Group. The official merger of Eurex Germany and the Swiss derivatives market Eurex Zürich (formerly SOFFEX) has created a genuine European futures and options market. Eurex is run on a joint electronic trading platform utilising harmonised trading guideline and admission procedures as well as a common clearing house.

of option prices in the GH model for a given volatility, e.g. a volatility estimated from time series data.

The rescaling of GH distributions is based on the shape- and location-invariant parameters given in Lemma 1.5. A consequence of this lemma is that the variance of the generalized hyperbolic distribution has a linear structure $\mathrm{Var}\,X_1 = \delta^2 C_\zeta$, where $C_\zeta$ depends only on the shape, i.e. the scale- and location invariant parameters. Therefore, one could also use $\delta$ as a volatility parameter.

To rescale the distribution such that it equals a given variance $\widehat{\sigma}^2$, one obtains the new $\widetilde{\delta}$ by

$$\widetilde{\delta} = \widehat{\sigma}\left(\frac{\mathrm{K}_{\lambda+1}(\widehat{\zeta})}{\widehat{\zeta}\,\mathrm{K}_\lambda(\widehat{\zeta})} + \frac{\widehat{\beta}^2}{\widehat{\alpha}^2 - \widehat{\beta}^2}\left[\frac{\mathrm{K}_{\lambda+2}(\widehat{\zeta})}{\mathrm{K}_\lambda(\widehat{\zeta})} - \left(\frac{\mathrm{K}_{\lambda+1}(\widehat{\zeta})}{\mathrm{K}_\lambda(\widehat{\zeta})}\right)^2\right]\right)^{-1/2}, \tag{2.1}$$

where $\left(\widehat{\alpha},\widehat{\beta},\widehat{\delta}\right)$ and $\widehat{\zeta}$ are estimated from a longer time series. With Lemma 1.5 the term in the brackets is also scale- and location invariant. To fix the shape of the distribution while rescaling with a new $\widetilde{\delta}$, one has to change the other parameters of the first parametrization in the following way

$$\widetilde{\lambda} = \widehat{\lambda}, \quad \widetilde{\alpha} = \frac{\widehat{\alpha}\,\widehat{\delta}}{\widetilde{\delta}}, \quad \widetilde{\beta} = \frac{\widehat{\beta}\,\widehat{\delta}}{\widetilde{\delta}} \text{ and } \widetilde{\mu} = \widehat{\mu}. \tag{2.2}$$

In the sequel we apply long-term estimates of the shape for the German stock options, estimated from daily time series from January 1, 1988 to May 24, 1994, and often rescale the generalized hyperbolic distribution for a short-term volatility estimate. Note, that by changing to the fourth parametrization the second step (2.2) is no longer necessary. The densities and moments in the forth parametrization are given in Appendix C.

## Visualisation

Analysing intraday option data sets, it is nearly impossible to see characteristic structures in the numerical results for each quote or single series. Graphics are much more comprehensible and allow a fast and frequent control of the results especially during the application of new mathematical or statistical approaches. An appropriate smoothing reduces the noise and the structure in the data becomes clearer. On the one hand tables have the advantage that they provide the exact numerical figures, but we believe that the patterns in the option pricing behaviour and implicit volatilities become more obvious with an adequate visualisation.

The curves in the plots are smoothed by the LOESS algorithm (Cleveland, Grosse, and Shyu 1993). In this approach it is only assumed that the parameter could be fitted locally by a polynomial of first or second order. We frequently do not want to make global assumptions regarding parameters in option data sets. Hence, we adapted the S tools for the LOESS smoothing algorithm to option data sets and modified it to prevent unnecessary extrapolation.
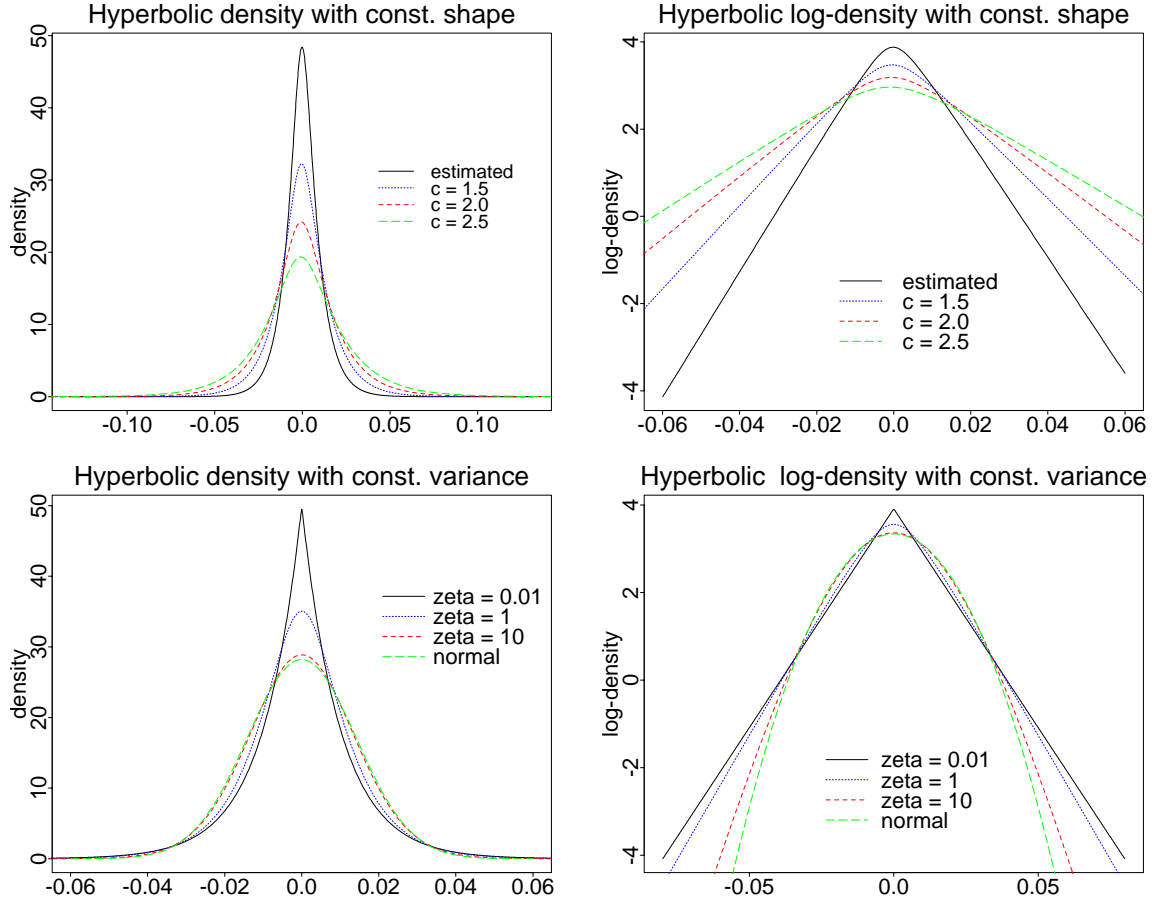
Figure 2.1: *Rescaled hyperbolic density (parameters estimates from Bayer stocks ($\delta' = c\,\widehat{\delta}$, $\zeta = 0.61$, $\varrho = 0.039$) and hyperbolic densities with const. variance and different shapes.*

## 2.2 Model versus Underlying Process

The first step in the evaluation of a particular option pricing model is the comparison of the model for the price process with observed prices. Samuelson (1965) proposed to model prices with the geometric or in his words "economic" Brownian motion. The famous Black and Scholes model is based on this reasoning. Nevertheless, already in the 60's began the search for more realistic models.

Eberlein and Keller (1995) proposed to model daily returns with the hyperbolic distribution and provide an extensive statistical examination. We discussed the fit of the hyperbolic distribution in the context of the extension of the hyperbolic model to generalized hyperbolic distributions in Chapter 1. The GH distribution provides also an excellent fit to the marginal distribution of intraday-returns (see Section 1.5 for intraday FX data). See Eberlein, Keller, and Prause (1998) for an overview over alternative modelling approaches.

### Dependence Structure

Typically, we find a nonsignificant correlation of the log-returns of most financial assets and a significant correlation of the squared returns at a confidence level of 5%. As an example we
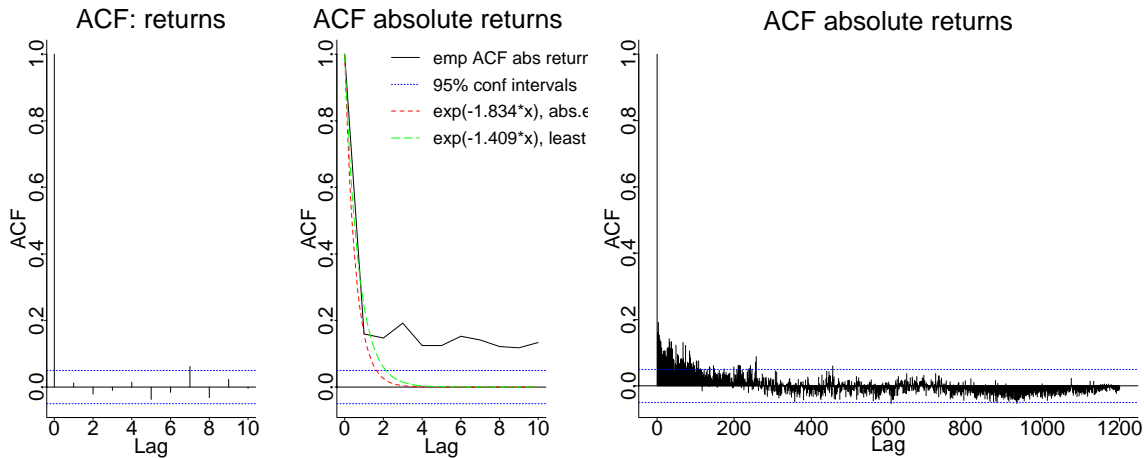
Figure 2.2: *Empiciral autocorrelation functions (ACF) of the returns and the absolute returns of Deutsche Bank stock from 1988–94 with 95% confidence interval.*

show the empirical autocorrelations of Deutsche Bank returns in Figure 2.2. The low correlation of log-returns hints at the efficiency of stock markets and also explains the difficulties to predict the future value of a stock price only from time series data. The correlation of squared returns has led to the development of GARCH and stochastic volatility models. Both models, the discrete and the continuous one rely on the fact that not the direction of the change but the magnitude depends on past behaviour in the share value.

The GH model is based on a Lévy motion, i.e. it is assumed that the returns are independent. This implies that squared returns are uncorrelated. Therefore, the introduction of stochastic volatility may improve the generalized hyperbolic model.

It is indeed difficult to handle temporal dependence in diffusion models. As an alternative Barndorff-Nielsen (1998) proposed a model in which the stochastic volatility is given by an Ornstein-Uhlenbeck type process. He replaced the GIG distribution in the mixing representation (1.6) by an Ornstein-Uhlenbeck-type process which is driven by a background driving Lévy process. See Chapter 3 for details. Finally, the construction led to a process with an ACF of the form $\exp(-\lambda u)$. We fitted this function to the empirical ACF by a minimization of the absolute and the squared errors. From the middle plot in Figure 2.2 it is clear, that it is not possible to model the structure of the empirical ACF exactly by the means of this simple model. Nevertheless, it is possible to refine this model by replacing the inverse Gaussian Ornstein-Uhlenbeck type (IGOU) process by a superposition, i.e. a finite sum of independent IGOU processes (Barndorff-Nielsen, Jensen, and Sørensen 1995, Chapter 3). Barndorff-Nielsen (1998) obtained promising results modelling the ACF of Deutsche Bank absolute returns by a superposition of two IGOU processes. Using this approach has the advantage that it leads to statistically and analytically tractable models with stochastic volatility, GH marginal distributions, and an analytically given ACF.

Another approach based on GH distributions without independent returns is given by Rydberg (1999). She proposes a diffusion model for stock returns, following Bibby and Sørensen (1997), where the stationary distribution is GH. However, neither the increments nor the ACF seem to be modelled appropriately by fitting the stationary distribution of an ergodic diffusion process. It is not clear, if the proposed process is close to observed processes.
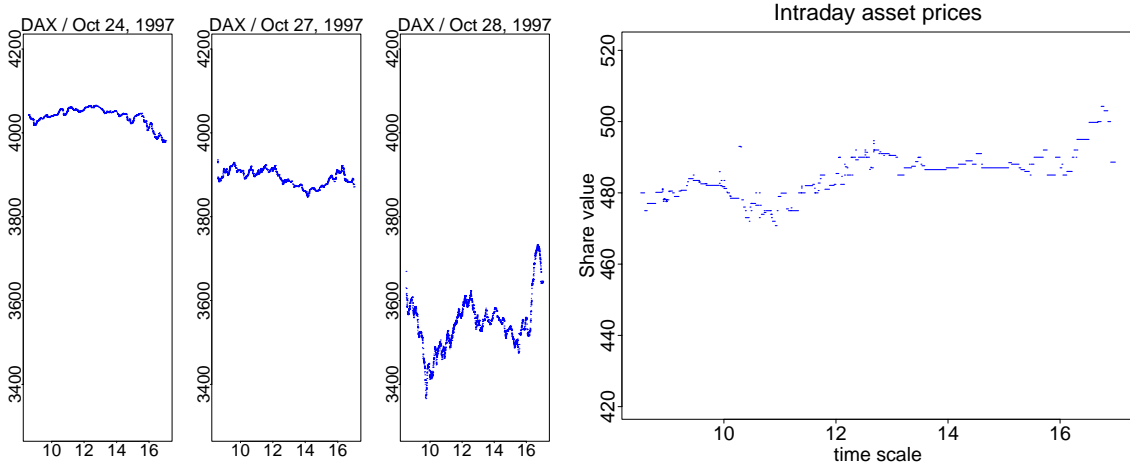
Figure 2.3: *Intraday value of the DAX during the crash 1997 (October 24-28, 1997) and Mannesmann October 28, 1997.*

**Scaling Laws**

One important aspect in the modelling of financial assets is the choice of the time scale. Analyzing price processes on different scales indicates the use of different models. Moreover, using wavelet transforms Cont (1999) has shown that often no existing model fits to the observed frequency structure on *every* scale.

Intraday behaviour of price processes is characterized by frequent jumps (see Figure 2.3). Thus, the hyperbolic Lévy motion as a pure jump process provides a better description of the stock market microstructure than processes with continuous paths. Another "stylized fact" of the intraday behaviour of financial assets are daily and weekly seasonalities (Guillaume, Dacorogna, Davé, Müller, Olsen, and Pictet 1997). This more complicated structure is usually modelled following a time series approach. See Kallsen and Taqqu (1998) for an option pricing approach in GARCH models. In Section 3.14 we propose a model which exhibits volatility clustering and has a purely discontinuous price process.

Barndorff-Nielsen and Prause (1999) hint at an explanation of the scaling laws of absolute returns typically observed in the foreign exchange market by NIG Lévy processes.

## 2.3 Options Sensitivities

After the comparison of the GH Lévy motion as a model for returns of financial assets with some stylized facts of observed price processes, we look closer at the properties of GH prices obtained by Esscher transforms. In this section we investigate the role of the different parameters for the pricing of options. This simulation study is based on estimated values of GH distributions for German stocks (but not on data sets of quoted options). The use of the Black-Scholes model as a benchmark is only a first step to understand a new model. Later on we confront the theoretical models with reality, i.e. intraday data from secondary markets.

At first, we compute the difference of Black-Scholes minus generalized hyperbolic respectively hyperbolic and NIG prices. The typical pattern of these differences is given in Figure 2.4. For options with a short time to maturity we observe the W-shape which was already

observed by Eberlein and Keller (1995) for hyperbolic prices. For the approximately symmetric GH distributions usually estimated for financial data, the price difference decreases with time-to-maturity, but it does not vanish for the usual range of maturities. Also for long-term options Black-Scholes prices at-the-money are higher than GH prices.

The W-shape is a result of the higher kurtosis of risk-neutral GH distributions. Consequently, we obtain a more pronounced W-shape for GH and NIG distributions than hyperbolic distributions, since they have more mass in the tails. Also for longer maturities the difference to Black-Scholes prices is bigger in the GH and NIG case. Recall, that the NIG prices are computed without Fast Fourier Transform. This allows to minimize the risk of numerical errors, since we can compare the prices in both approaches. Although it is impossible to exclude every numerical error, there are at the worst only small numerical errors in the subsequent computations left.

We summarize results in a short overview below. Note, that the simulation results are only valid for the GH distributions usually estimated from financial time series data. We have plotted the difference of hyperbolic minus Black-Scholes prices on page 45 for of hyperbolic distributions (we look at subclasses because this allows us to use the shape triangle in Figure 2.5 to compare the shape parameters of different stocks). However, the results are similar for other GH distributions. For instance, compare the results for the hyperbolic model on page 45 with those for the NIG model on page 46.

**Kurtosis:** The GH distribution is more peaked in comparison to the normal distribution. This leads to a W-shape in the price differences for short maturities (see Figure 3.16 in Keller 1997), and allows for a correction of the smile effect of implicit volatilities (cf. Section 2.4).

In Figure 2.5 and 2.8 we have plotted the price difference for some hyperbolic and NIG distributions with different shapes. To eliminate the effect of the varying volatilities, we have rescaled all hyperbolic distributions to an annual volatility of 20%. The shape parameters are given in the triangle on the right side: The Deutsche Bank (DBK) estimate has a substantially higher kurtosis, therefore, the W-shape is more emphasized than the W-shape for Bayer and Thyssen calls.

**Time to Maturity:** The difference of GH prices and Black-Scholes prices decreases with time to maturity (see page 44). This is a result of the aggregational Gaussianity in the generalized hyperbolic model.

On the other hand, if we assume a very high skewness, the difference of GH minus Black-Scholes price increases with time to maturity (see Figure 2.6).

**Skewness:** In Figure 2.6 and 2.9 we have shown the effect of the skewness (given by $\chi$) on the price difference. Note, that the values of $\chi$ are larger than usually observed for the hyperbolic distribution.
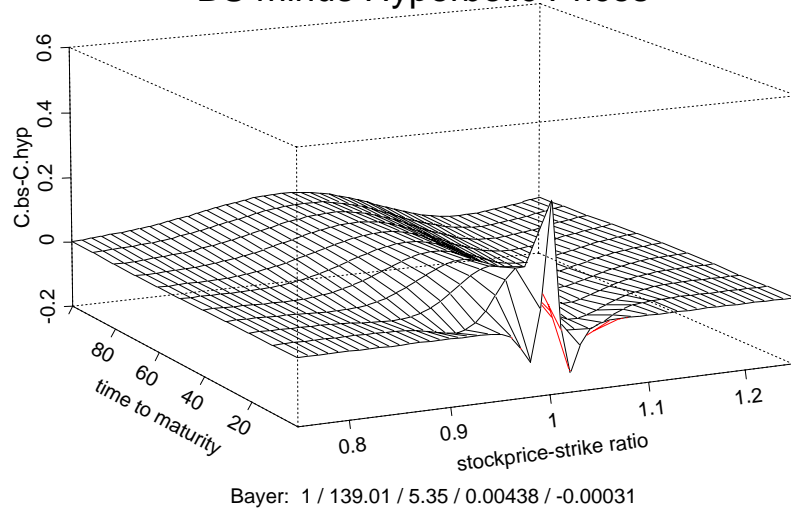
The effect of skewness depends in the following way on the construction of the simulation study: we have assumed that the volatility of the real-world measures is the same in the Black-Scholes and the GH model. The Esscher transform leads to a smaller volatility in the Esscher martingale measure (see Figure 2.11).

Only the absolute magnitude of skewness matters, whereas the sign of the skewness parameter $\chi$ has no influence on the price in the GH model: the curves in Figure 2.6 are

43

**BS minus Generalized Hyperbolic Prices**

Bayer: -1.79 / 21.34 / 2.67 / 0.01525 / -0
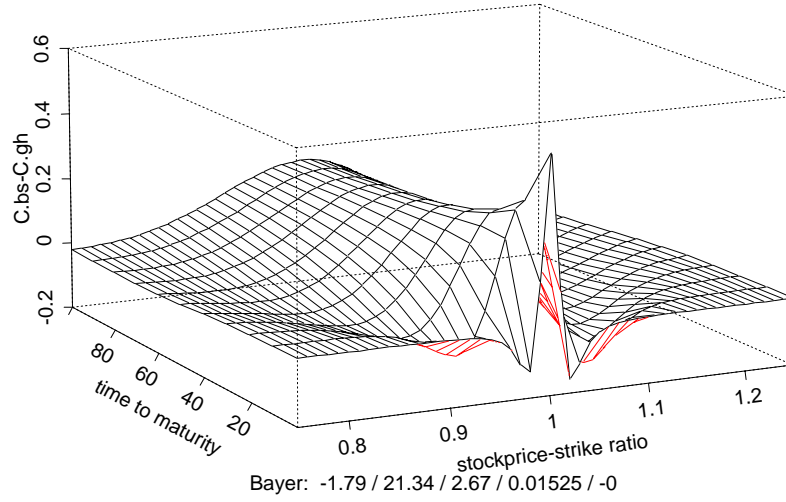
**BS minus Hyperbolic Prices**

Bayer: 1 / 139.01 / 5.35 / 0.00438 / -0.00031

**BS minus NIG Prices**

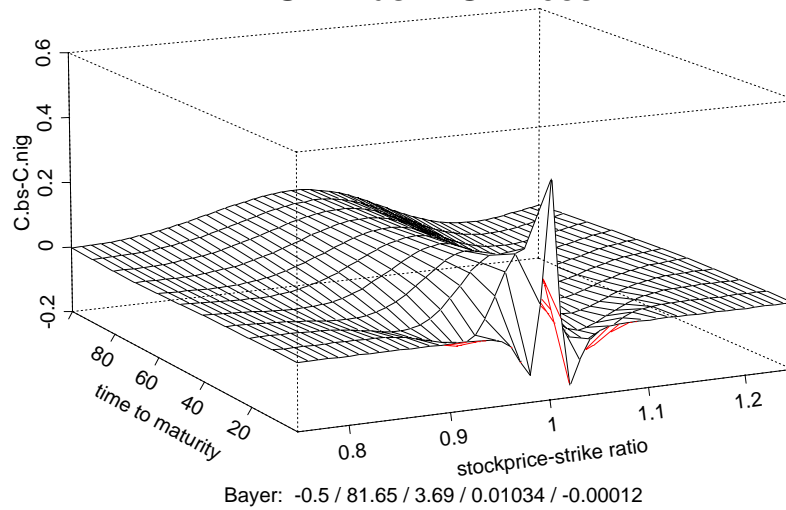Bayer: -0.5 / 81.65 / 3.69 / 0.01034 / -0.00012

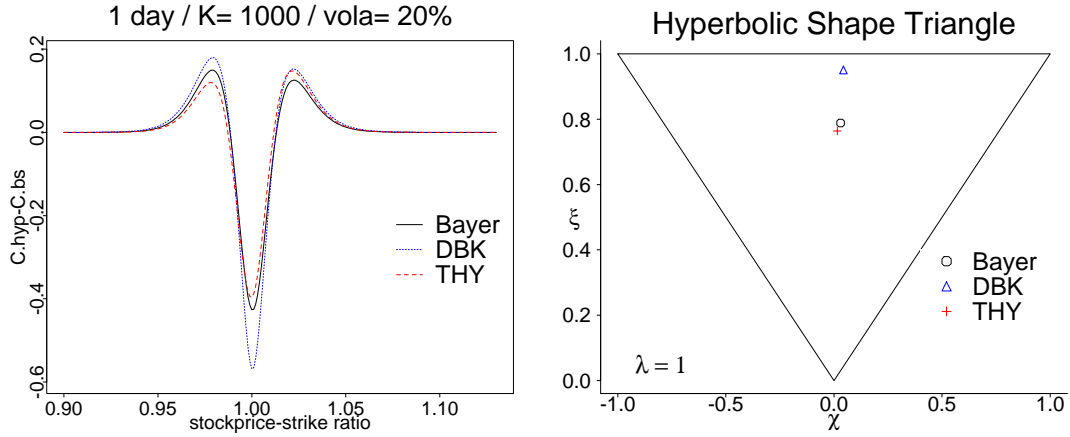Figure 2.4: *Black-Scholes minus GH prices (Bayer parameters, strike = 1000).*

Figure 2.5: *Hyperbolic minus Black-Scholes prices for different hyperbolic distributions (rescaled to a volatility of 20%, interest rate 0%. The likelihood estimates in the shape triangle are given in Table 1.14).*
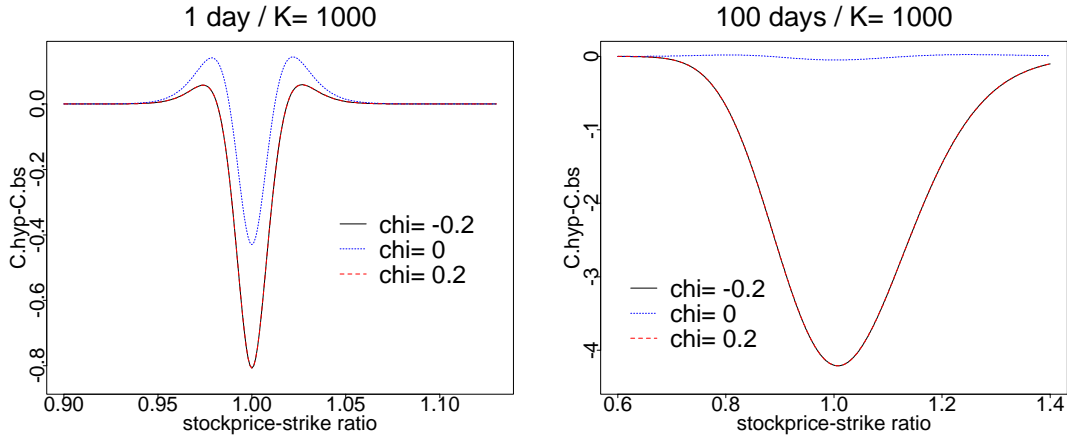


Figure 2.6: *Sensitivity w.r.t. skewness: Hyperbolic minus Black-Scholes prices. The price differences for $\chi = -0.2$ and $\chi = 0.2$ are identical (annual volatility 20%, interest rate 0%, $\xi = 0.8$ and $\mu = 0$).*
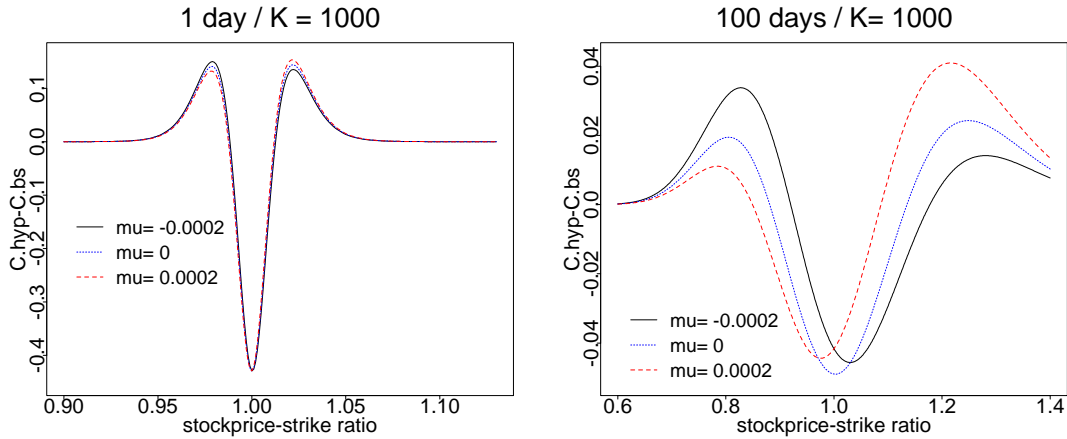


Figure 2.7: *Sensitivity w.r.t. drift: Hyperbolic minus Black-Scholes prices (annual volatility 20%, interest rate 0%, $\xi = 0.8$ and $\chi = 0$).*
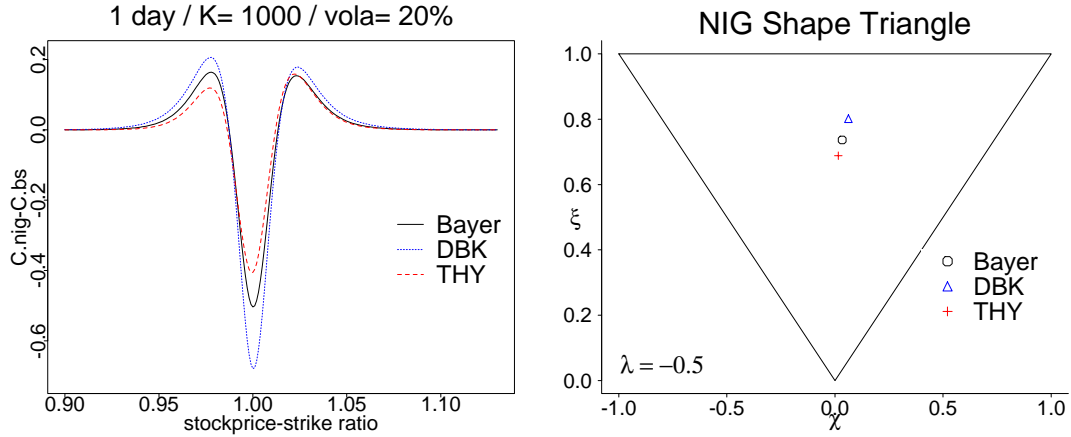
Figure 2.8: *NIG minus Black-Scholes prices for different NIG distributions (rescaled to a volatility of 20%, interest rate 0%, the likelihood estimates are given in Table 1.14).*
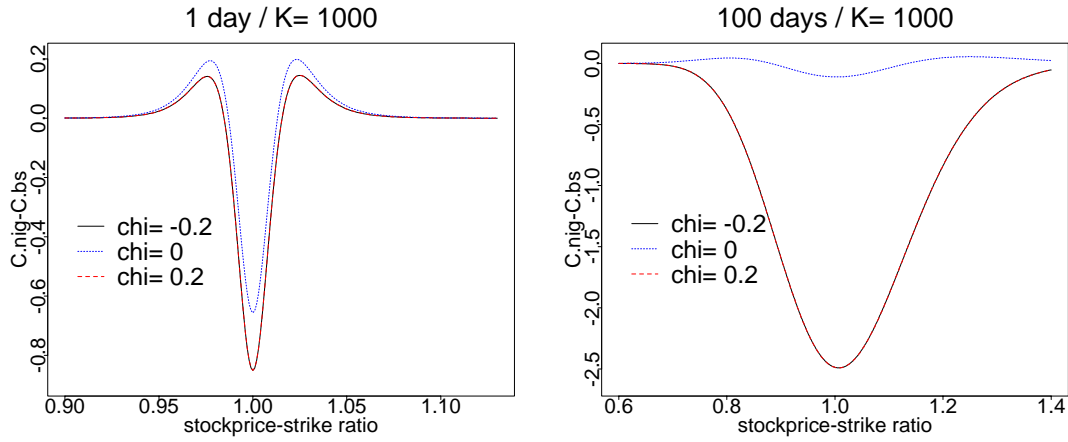


Figure 2.9: *Sensitivity w.r.t. skewness: NIG minus Black-Scholes prices. The price differences for $\chi = -0.2$ and $\chi = 0.2$ are the same (annual volatility 20%, interest rate 0%, $\xi = 0.8$ and $\mu = 0$).*
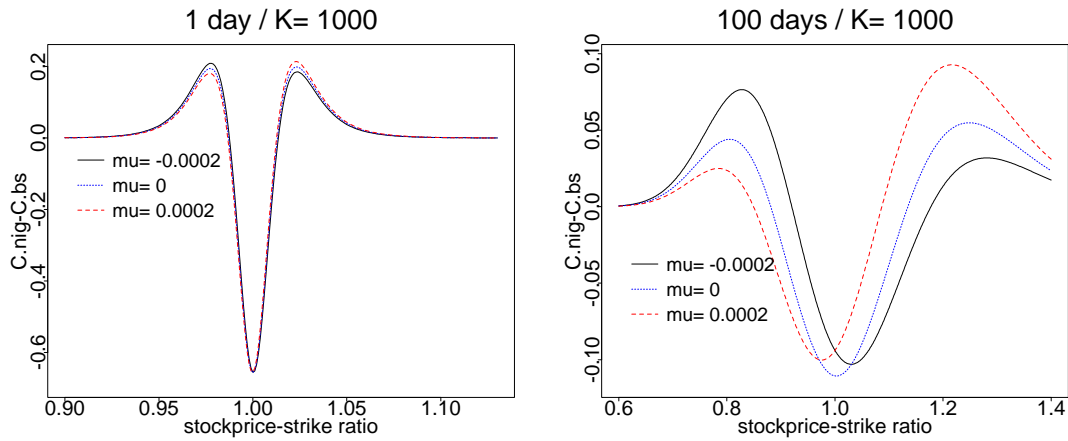


Figure 2.10: *Sensitivity w.r.t. drift: NIG minus Black-Scholes prices (annual volatility 20%, interest rate 0%, $\xi = 0.8$ and $\chi = 0$).*
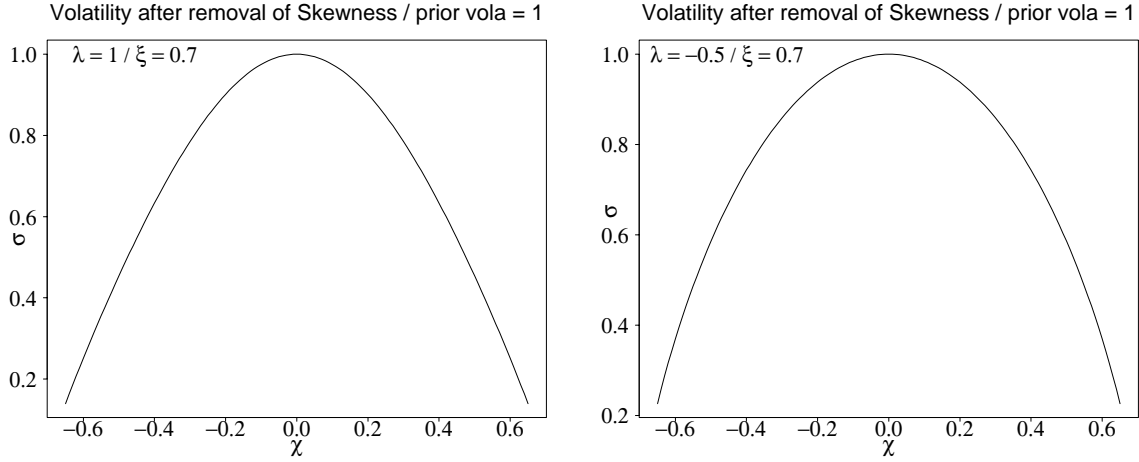
46

Figure 2.11: *Volatility of a GH distribution after the removal of the skewness.*

identical. Note, that the variance in (1.16) depends only on the square of the skewness parameter $\beta$. This explains why the curves are identical.

On the other hand, the volatility remains unchanged by the Girsanov transformation in the Black-Scholes model. Hence, GH prices have to be smaller. If we compare prices in the GH and the Black-Scholes model with an equal variance under the risk-neutral measure, the difference vanishes as a result of the aggregational Gaussianity of the GH Lévy motion. See Raible (1999) for a comparison of prices based on an equal volatility under the risk-neutral measure.

**Drift:** In the Black-Scholes model the drift is removed by the Girsanov transformation. Hence, it plays no role for the pricing of options. The Esscher transformation skews the exponential Lévy process until it is a martingale, but the exponential tilting does not remove the drift itself. Therefore, the choice of the drift has a small influence on the price. The simulation results in Figure 2.7 show that the effect of a drift decreases when the time to maturity decreases.

In particular a positive skewness leads to an increased difference of GH minus Black-Scholes prices in-the-money and vice versa.

## 2.4 Implicit Volatilities

Writing the actual market price on the left side of the equation and the Black-Scholes option pricing formula on the right side and solving for the volatility parameter $\sigma$ yields the Black-Scholes implicit volatility $\sigma_{\text{Imp.BS}}$. This is the volatility assumed by the traders. According to both models the volatility should be constant for different stockprice-strike ratios $\varrho = S^-/K$. Figure 2.12 shows for the Black-Scholes model, that in reality $\sigma_{\text{Imp.BS}}$ depends on the stockprice-strike ratio and on time to maturity.

Typically the implicit volatility is higher in-the-money and out-of-the-money. This effect is called the smile because the shape of the curve resembles a smiling face. The smile is decreasing with time to maturity and has its minimum for $\varrho = S^-/K \approx 1$. Looking at data
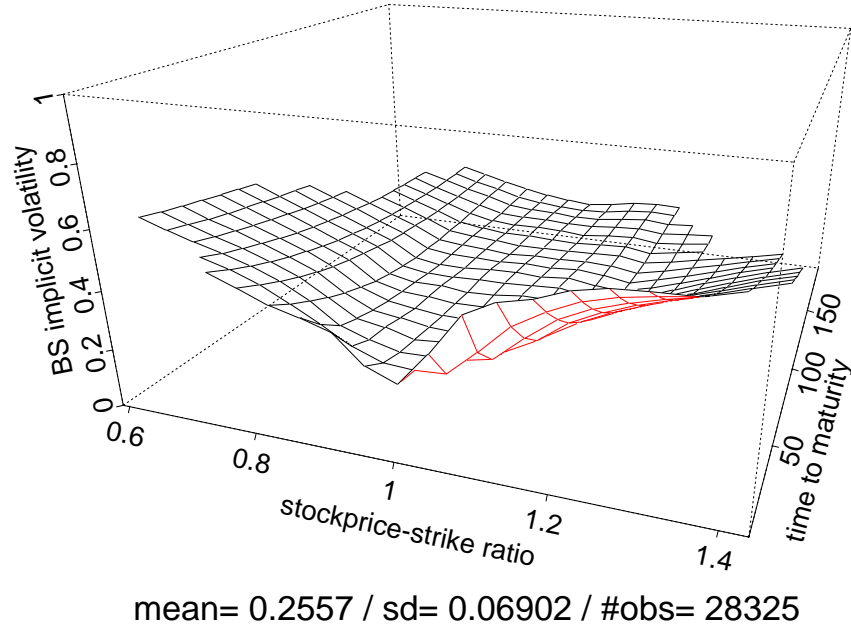
mean= 0.2557 / sd= 0.06902 / #obs= 28325

Figure 2.12: *Black-Scholes implicit volatilities of Thyssen calls from July 1, 1992 to August 18, 1994.*

sets corresponding to different time periods, the implicit volatility follows always the same pattern but of course the smile is more regular for longer observation periods. The pattern repeats for all analysed share values and for the option on the DAX future given in Figure 2.13. Below the interdependence of implicit volatility and stockprice-strike ratio respectively time to maturity is plotted separately for Thyssen calls.

But the "smirks" of implicit volatilities that were reported for the S&P future and index options (see also Derman and Kani 1994, and Longstaff 1995) show a different pattern than the smile effect we observed for individual German stock and DAX future options. We also looked at the implicit volatilities of options on the DAX in the year 1996. There we observed an asymmetric smile in comparison to the options on stocks and for options on futures in the German market. The liquidity of the DAX options is much higher than the liquidity of single stock options, therefore, the (symmetric) effect of a low liquidity of deep-in- or deep-out-of-the-money options may be reduced.

### Reduction of the Smile

One of the first who systematically and empirically studies alternative but now outdated option pricing formulas was Rubinstein (1985). None of the models he examined correct all the observed deficiencies of the Black-Scholes model. Therefore, he proposed to build a composite model, or to correlate the bias of the option prices to macroeconomic variables.

Eberlein, Keller, and Prause (1998) give an overview over various approaches to reduce the smile effect. It is worth to mention, that most pricing models recently proposed lead theoretically, i.e. in simulation studies, to a reduction of the smile. On the contrary, empirical studies with a representative data set from the secondary market are seldomly available.
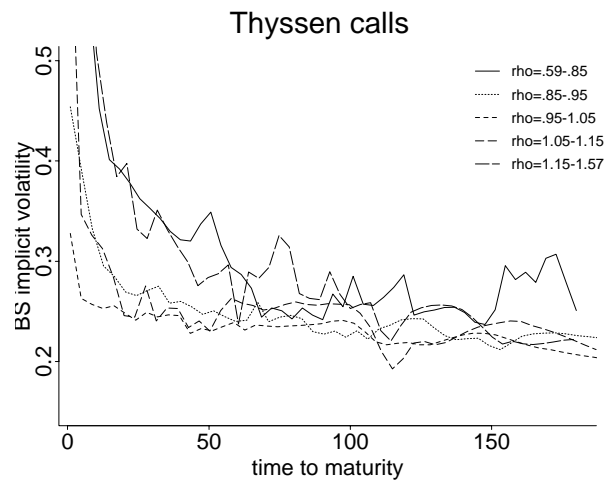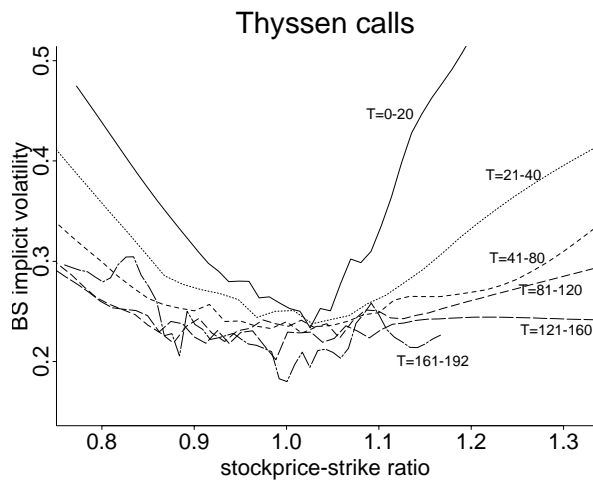
## Deutsche Bank calls

mean= 0.1879 / sd= 0.065 / #obs= 68803

## Siemens calls

mean= 0.1792 / sd= 0.054 / #obs= 78571

## DAX future calls  7-12.1994

mean= 0.1876 / sd= 0.032 / #obs= 830

## DAX future puts  7-12.1994

mean= 0.2186 / sd= 0.02978 / #obs= 467

## Thyssen calls

T=0-20
T=21-40
T=41-80
T=81-120
T=121-160
T=161-192

## Thyssen calls

rho=.59-.85
rho=.85-.95
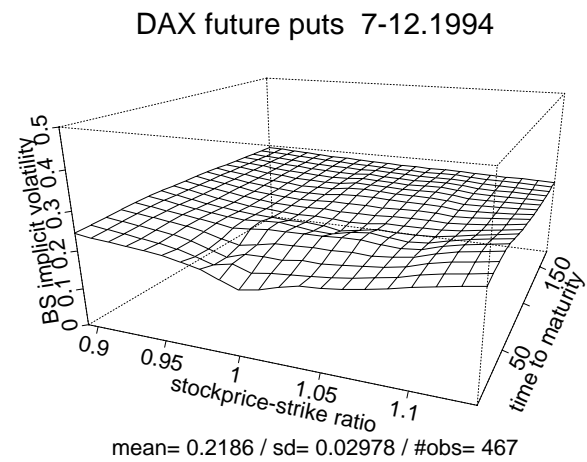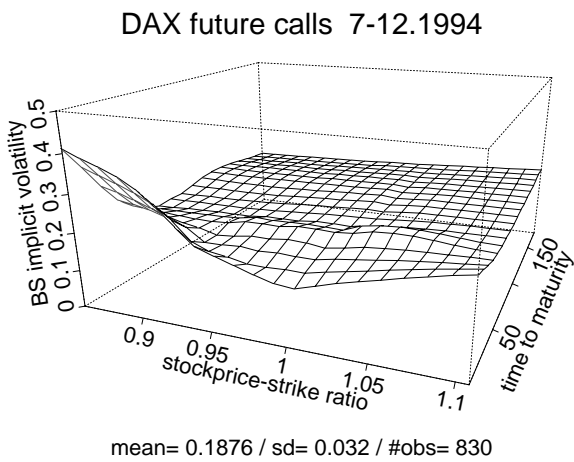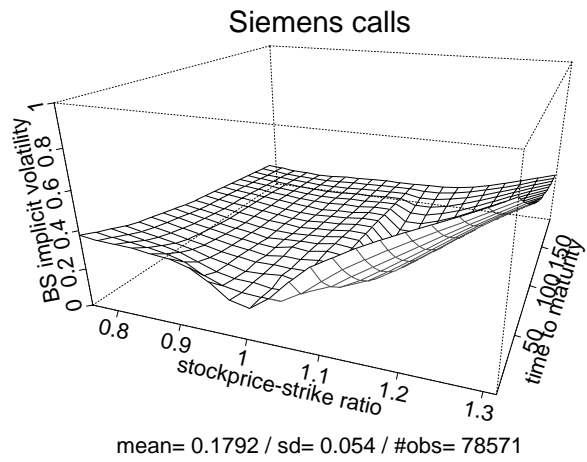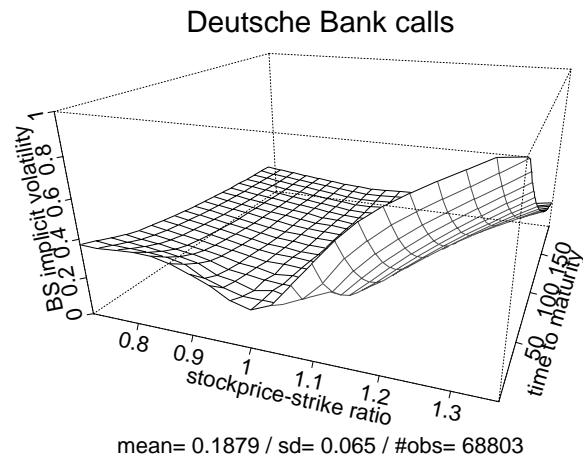rho=.95-1.05
rho=1.05-1.15
rho=1.15-1.57

Figure 2.13: *Black-Scholes implicit volatilities of DTB stock options from July 1, 1992 to August 19, 1994 and of DAX futures options from July 1, 1994 to December 29, 1994.*

49

## Smile Effect in the Generalized Hyperbolic Model

Now we examine to which extend the replacement of the Gaussian model by the generalized hyperbolic one leads to a better option pricing behaviour. For the comparison of the Black-Scholes and the GH model we use GH parameters derived from the variance by the rescaling mechanism described in (2.1) and (2.2). The implicit generalized hyperbolic volatility $\sigma_{\text{Imp.GH}}$ is computed in the same way as in the Black-Scholes case.

The empirically observed $\sigma_{\text{Imp.Hyp}}$ is shown in Figure 2.14 (top left) for Deutsche Bank call options. The implicit Black-Scholes volatility for the same data set is given in Figure 2.13 (top left). At first sight the implicit hyperbolic volatility produces a smile effect similar to the one arising in the Black-Scholes setting. Plotting the difference $\sigma_{\text{Imp.BS}} - \sigma_{\text{Imp.Hyp}}$ of implicit volatilities of the two models (top right) shows that in the hyperbolic case the smile effect is reduced according to the W-shape. In the second row we plotted the difference of Black-Scholes minus NIG resp. GH implicit volatility. Obviously, we see a larger correction of the smile effect in the NIG and the largest correction in the GH model. The two remaining plots in Figure 2.14 show the difference $\sigma_{\text{Imp.BS}} - \sigma_{\text{Imp.Hyp}}$ as a function of moneyness and as a function of time to maturity respectively.

It is worth to mention that the differences $\sigma_{\text{Imp.BS}} - \sigma_{\text{Imp.Hyp}}$ of the implicit volatilities are not symmetric around $\varrho = S^-/K = 1$. Consequently we also observe a correction of the asymmetries in the smile—the "smirk" effect—which is only small in the German stock options datasets.

Another way to analyse the smiles in both models is to fit the following linear model for the implicit volatilities

$$\sigma_{\text{Imp},i} = b_0 + b_1 T_i + b_2 \frac{(\varrho_i - 1)^2}{T_i} + e_i, \qquad (2.3)$$

where $e_i$ is the random error term and $i$ the number of the trade in the option data set. The cross-term $(\varrho - 1)^2/T$ reflects the degression of the smile effect with increasing time to maturity $T$. The regression function was chosen as parsimonious as possible. Results of this regression are given in Table 2.15. Note, that $T$ is measured in trading days and therefore, the time-to-maturity effect is of a relevant order. As the value of the coefficients for $(\varrho - 1)^2/T$ and for $T$ are smaller for the hyperbolic, the NIG and the GH model, we conclude that these models reduce the smile effect. As we have already seen in Figure 2.14, we observe the largest correction of the smile effect in the GH model. We provide in Table 2.15 also the results for the symmetric centered versions of the models which yield option prices similar to the Black-Scholes model for longer maturities. Therefore, we do not observe any reduction of the time-to-maturity effect but the expected decrease of the smile.
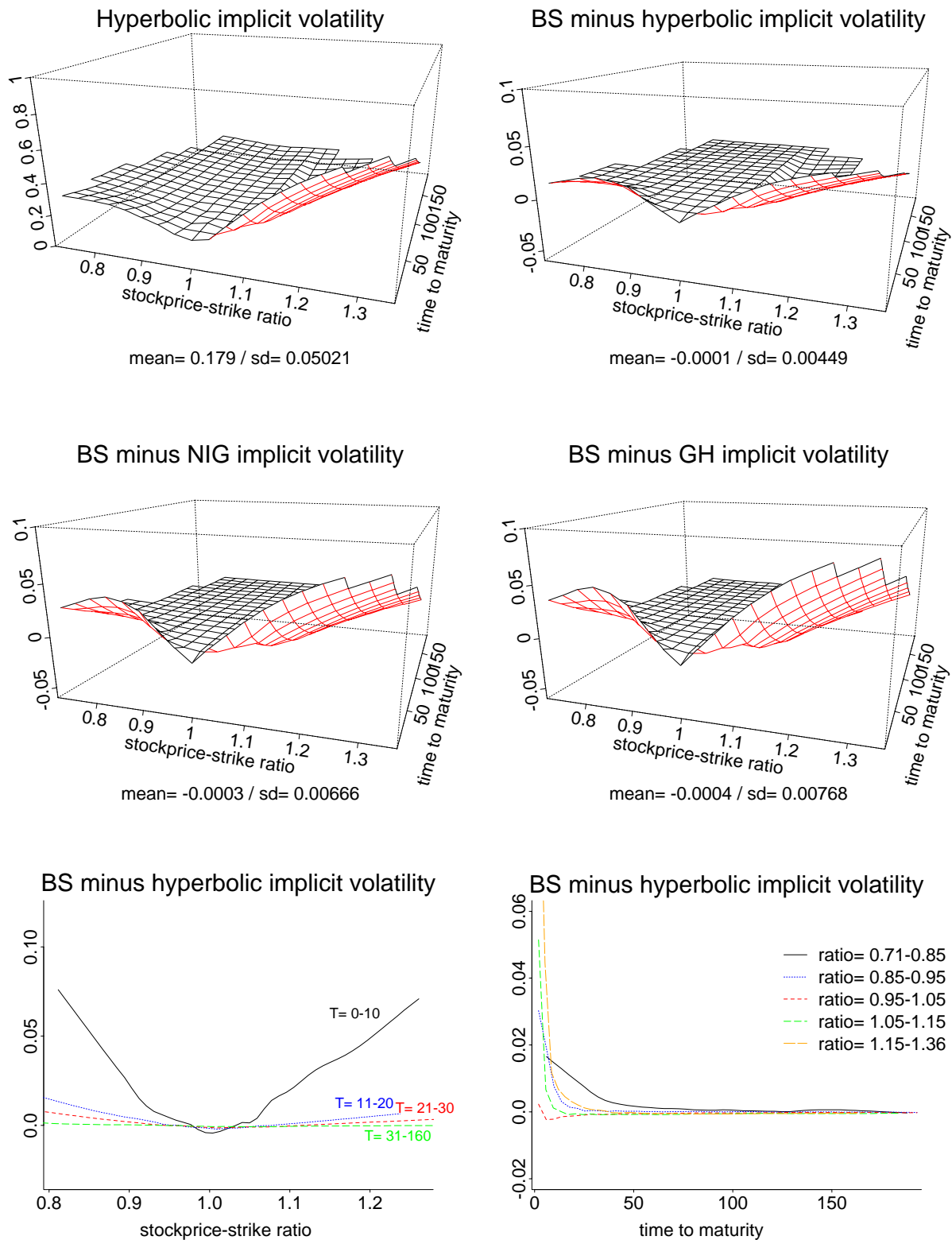
Figure 2.14: *Implicit hyperbolic volatility and comparison of the implicit volatilities of the Black-Scholes, the hyperbolic the NIG and the generalized hyperbolic model (Deutsche Bank calls from July 1992 to August 1994, 68803 obs.).*

Table 2.15: *Fitted coefficients for call options from July 1992 to August 1994. "SC" marks the results for the symmetric centered versions of the models.*

|  |  | $b_0$ | $b_1$ | $b_2$ | $R^2$ |
|---|---|---|---|---|---|
| Bayer | Black-Scholes | 0.1901 | $-0.00033$ | 51.77 | 0.5432 |
|  | Hyperbolic | 0.1907 | $-0.000334$ | 48.66 | 0.516 |
|  | HYP SC | 0.1904 | $-0.000332$ | 48.44 | 0.5156 |
|  | NIG | 0.1909 | $-0.000334$ | 47.45 | 0.5057 |
|  | NIG SC | 0.1906 | $-0.000333$ | 47.31 | 0.5055 |
|  | GH | 0.1915 | $-0.000337$ | 46.22 | 0.4932 |
|  | GH SC | 0.1911 | $-0.000336$ | 46.13 | 0.4932 |
| Daimler Benz | Black-Scholes | 0.2177 | $-0.00029$ | 40.53 | 0.5416 |
|  | Hyperbolic | 0.2186 | $-0.0003$ | 36.89 | 0.4972 |
|  | HYP SC | 0.2184 | $-0.000293$ | 36.33 | 0.4951 |
|  | NIG | 0.2191 | $-0.000305$ | 35.11 | 0.4746 |
|  | NIG SC | 0.2189 | $-0.000296$ | 34.48 | 0.4716 |
|  | GH | 0.2207 | $-0.000321$ | 32.81 | 0.4378 |
|  | GH SC | 0.2201 | $-0.000306$ | 31.98 | 0.4343 |
| Deutsche Bank | Black-Scholes | 0.1831 | $-0.000266$ | 57.75 | 0.4409 |
|  | Hyperbolic | 0.1843 | $-0.000273$ | 51.00 | 0.385 |
|  | HYP SC | 0.1840 | $-0.000271$ | 50.65 | 0.3836 |
|  | NIG | 0.1851 | $-0.000278$ | 47.71 | 0.3563 |
|  | NIG SC | 0.1846 | $-0.000275$ | 47.25 | 0.3543 |
|  | GH | 0.1856 | $-0.000281$ | 46.16 | 0.3426 |
|  | GH SC | 0.1850 | $-0.000278$ | 45.73 | 0.3408 |
| Siemens | Black-Scholes | 0.1712 | $-9.7e-05$ | 86.92 | 0.4673 |
|  | Hyperbolic | 0.1728 | $-0.000106$ | 77.17 | 0.4043 |
|  | HYP SC | 0.1726 | $-0.000105$ | 76.83 | 0.4036 |
|  | NIG | 0.1735 | $-0.00011$ | 73.73 | 0.3805 |
|  | NIG SC | 0.1731 | $-0.000109$ | 73.34 | 0.3799 |
|  | GH | 0.1779 | $-0.000125$ | 69.67 | 0.3414 |
|  | GH SC | 0.1747 | $-0.000123$ | 68.34 | 0.3411 |
| Thyssen | Black-Scholes | 0.2504 | $-0.000214$ | 81.73 | 0.481 |
|  | Hyperbolic | 0.2522 | $-0.000225$ | 75.75 | 0.4422 |
|  | HYP SC | 0.2520 | $-0.000225$ | 75.94 | 0.4438 |
|  | NIG | 0.2525 | $-0.000227$ | 74.97 | 0.436 |
|  | NIG SC | 0.2524 | $-0.000228$ | 75.31 | 0.4386 |
|  | GH | 0.2522 | $-0.000225$ | 75.96 | 0.4438 |
|  | GH SC | 0.252 | $-0.000225$ | 76.05 | 0.4448 |
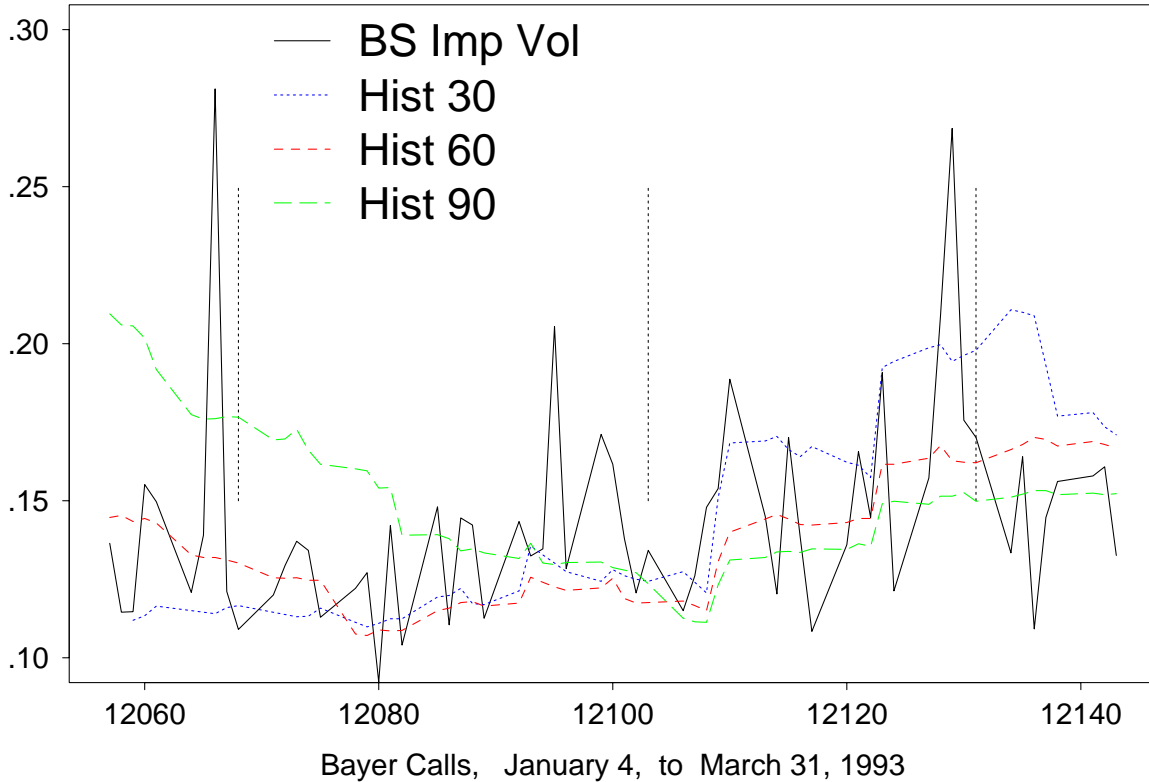
Figure 2.16: *Deficiencies of historical and implicit volatility estimators.*

## 2.5  Pricing Performance

An alternative approach to testing option pricing models is to compare observed market prices with model prices. In contrast to volatility comparisons pricing performance analyses price differences. Hence, the unit of measurement is a currency unit, in this case Deutsche Mark. Note, that the same difference in volatility has a greater effect on the price if the time to maturity is longer.

However, one remaining problem is to choose the volatility. We estimate volatility parameters following different approaches. First, we compute historical volatilities for time windows of 30 calendar days before the trading day of the option using the classical variance estimator. These are named Hist30 in the sequel. Secondly, we apply implicit volatilities observed before each trade. For the estimator Imp.median$n$ we take the running median of the implicit volatilities of the last $n$ quoted options. Because of its robustness the running median proved to be a better estimator than means or trimmed means. Note, that we follow out-of-sample approaches both for the historical and the implicit volatility. Cox and Rubinstein (1985) describe the option pricing service of Fisher Black, who used historical and implicit volatilities and some other market parameters for computing a volatility estimate. Thus, both procedures are used in practice.

Both estimation approaches for the volatility have their disadvantages: The implicit volatilities are peaked before the expiration day of the options. This can be seen in Fig-
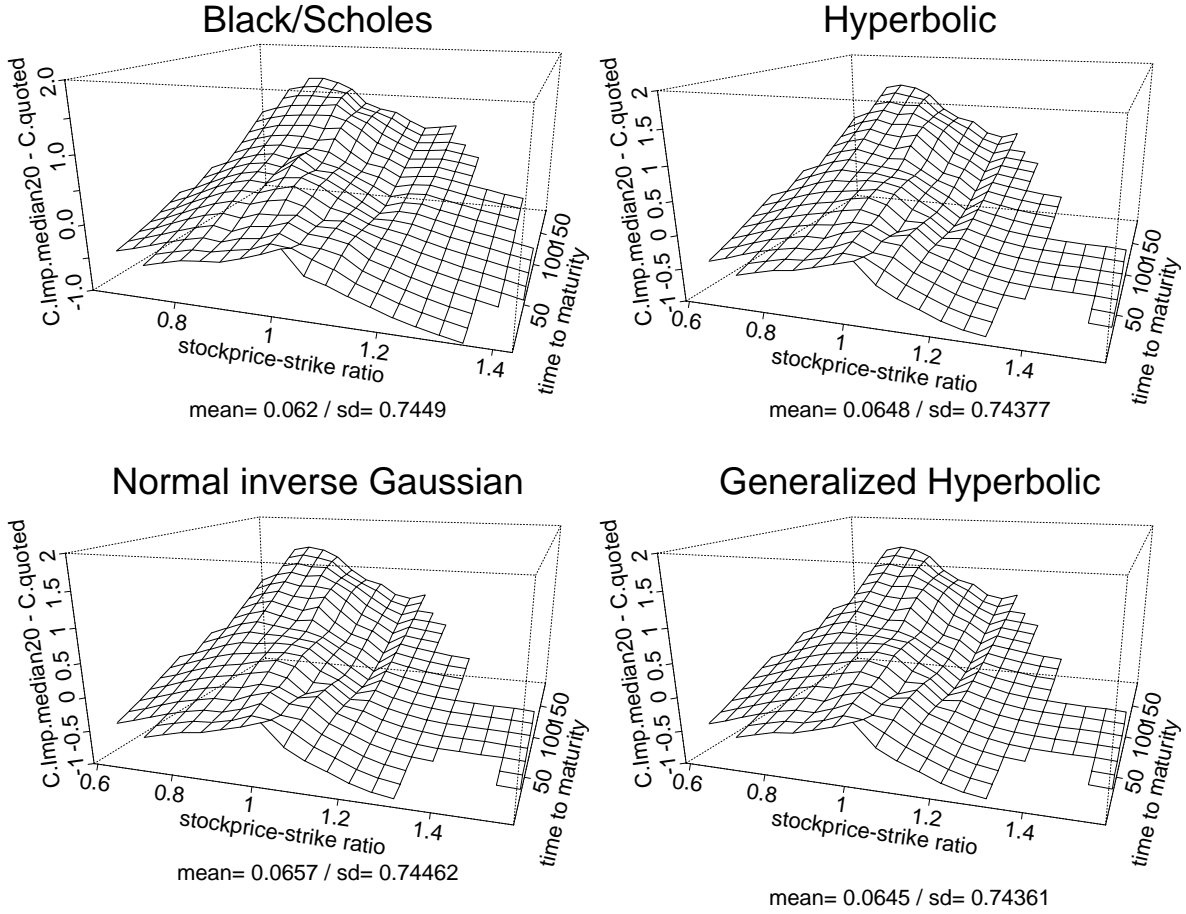
53

## Black/Scholes

C.Imp.median20 - C.quoted

stockprice-strike ratio

time to maturity

mean= 0.062 / sd= 0.7449

## Hyperbolic

C.Imp.median20 - C.quoted

stockprice-strike ratio

time to maturity

mean= 0.0648 / sd= 0.74377

## Normal inverse Gaussian

C.Imp.median20 - C.quoted

stockprice-strike ratio

time to maturity

mean= 0.0657 / sd= 0.74462

## Generalized Hyperbolic

C.Imp.median20 - C.quoted

stockprice-strike ratio

time to maturity

mean= 0.0645 / sd= 0.74361

Figure 2.17: *Pricing performance of Thyssen calls (July 1, 1992 to August 18, 1994, volatility estimator Imp.median20).*

ure 2.16, where the uninterrupted line shows the daily weighted average of the Black-Scholes implicit volatilies and the vertical dotted lines the expiration days at the DTB. The dashed and the dotted lines represent the historical volatility estimators. Obviously, they are dominated by single larger returns and their values are distinct for different window lengths. Of course it is possible to find more sophisticated versions of these estimators, e.g. variance estimators with decreasing weights or implicit volatility estimators using only quotes at-the-money, but the fundamental statistical problems do not change: First, the volatility is usually not assumed to be stationary. Therefore, only short time intervals should be used to estimated the instantaneous volatility. Furthermore, since the option price is not a linear function w.r.t volatility in pricing models, even an unbiased estimate of the "true" volatility may not yield unbiased estimates of true option prices. Moreover, even if the volatility estimate is unbiased, the variance of the estimates is high. Review Table 1.6 for the stability of the volatility estimation.

In Figure 2.17 we present typical plots of the pricing performance for the Black-Scholes and for the GH models. The difference of model price minus market price of the call options increases with time to maturity. We used the volatility estimator Imp.median20 which is based on implicit volatilities.
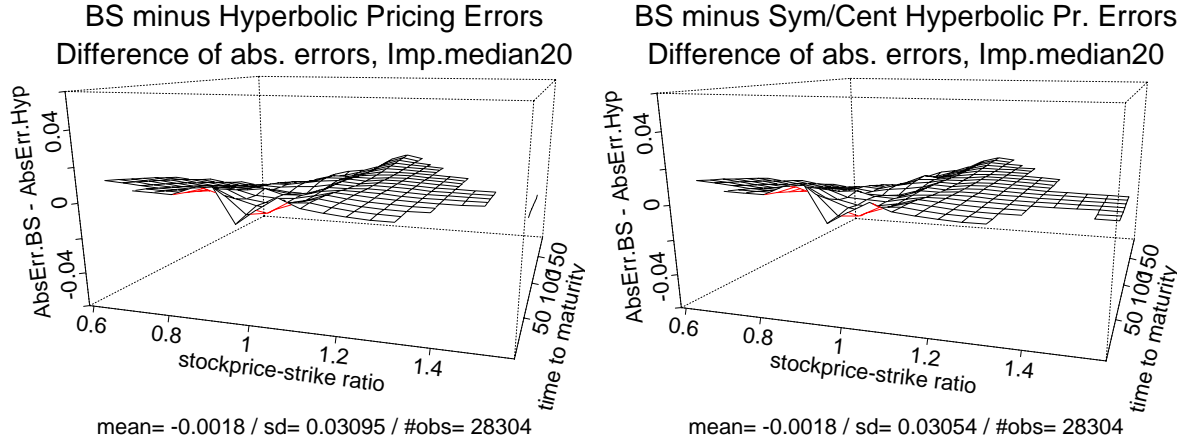
54

Figure 2.18: *Comparison of the pricing performance of Thyssen calls (July 1, 1992 to August 18, 1994, volatility estimator Imp.median20).*

Looking at the options with short maturities we see the inverse pattern to the smile of the implicit volatilities. All models lead to prices below the market price for short maturity options out-of-the-money or in-the-money, hence the implicit volatility surface becomes U-shaped.

A comparison of the pricing performance of the Black-Scholes and the hyperbolic model within a single plot is given in Figure 2.18 (left). We compute for each quote the difference of the absolute pricing errors of the two models: absolute pricing error using Black-Scholes minus absolute pricing error using the hyperbolic model. The plot reveals a distinct correction of the mispricing by the GH model for call options with short maturities according to the smile. In Figure 2.18 (right) we provide the same difference for the symmetric centered model. Obviously, the correction of pricing errors does primarily depend on the modelling of the kurtosis in the GH model. Skewness and drift are not equally important.

Considering the smile plots in Section 2.4 the deficiencies of the Black-Scholes model are worse for options near to expiration. The pricing error measured in Deutsche Mark is bigger for options with longer maturities. Consequently, we have to analyse both, pricing performance and smile, to get a complete picture for the two models.

Finally, we choose a global approach to compare the models. In Table 2.19 we give the mean (and the standard deviation) of the pricing errors for Bayer calls, i.e.

$$\frac{1}{N} \sum_{i=1}^{N} \left( C_{\text{model},i} - \widehat{C}_i \right), \tag{2.4}$$

where $\widehat{C}_i$ denotes the quoted option price of trade $i = 1, \ldots, N$.

It is obvious that prices based on implicit volatility estimators are closer to the market. In an efficient secondary market new information affecting the volatility of the underlying stock will be impounded in the option prices immediately. Clearly, estimation of volatilities from historical data will react more slowly. Nevertheless, Beckers (1981) observed that historical volatilities add information to implicit volatilites. Therefore, he concluded that at that time the secondary market was not informationally efficient. This has changed after the crash of

Table 2.19: *Pricing performance of pricing models w.r.t. different volatility estimators: mean and standard deviation of the difference of the model price minus quoted price (July 1, 1992 to August 19, 1994).*

| Estimator | Black-Scholes Mean | StDev | Hyperbolic Mean | StDev | NIG Mean | StDev | GH Mean | StDev |
|---|---|---|---|---|---|---|---|---|
| **Bayer** | | | | | | | | |
| Hist30 | 0.798 | 2.389 | 0.78 | 2.379 | 0.773 | 2.379 | 0.752 | 2.373 |
| Imp.median20 | 0.279 | 1.475 | 0.28 | 1.462 | 0.281 | 1.457 | 0.283 | 1.449 |
| Imp.median30 | 0.276 | 1.399 | 0.276 | 1.393 | 0.281 | 1.457 | 0.283 | 1.449 |
| **Thyssen** | | | | | | | | |
| Hist30 | 0.208 | 1.894 | 0.191 | 1.891 | 0.189 | 1.89 | 0.19 | 1.89 |
| Imp.median20 | 0.062 | 0.745 | 0.065 | 0.744 | 0.066 | 0.745 | 0.065 | 0.744 |
| Imp.median30 | 0.055 | 0.696 | 0.058 | 0.7 | 0.058 | 0.701 | 0.057 | 0.7 |

'87. See Christensen and Prabhala (1998) for a recent study of implied and realized volatility for S&P100 Index options. They especially emphasize the shift of regime during the crash '87, which led to a less noisy implied volatility in the post-crash period, and they find that implied volatility predicts future realized volatility as well as in conjunction with historical volatility estimates.

The construction of the estimator Imp.median$n$ allows on the one hand for the quick adaption to new information, on the other hand the estimator is robust enough against outliers in observed implicit volatilities. These outliers are likely to occur because prices change only in tick-size steps and small changes in the price do have a large impact on implicit volatilities for options with short maturities.

Another interesting aspect is to take trading volume into consideration. We compute the weighted mean of the difference of the absolute errors

$$\sum_{i=1}^{N} \left( |C_{\mathrm{BS},i} - \widehat{C}_i| - |C_{\mathrm{GH},i} - \widehat{C}_i| \right) \cdot \mathrm{Vol}_i \left/ \sum_{i=1}^{N} \mathrm{Vol}_i \right. , \tag{2.5}$$

where $\mathrm{Vol}_i$ denotes the volume of trade $i = 1, \ldots, N$ with quoted option price $\widehat{C}_i$. Furthermore, the model prices are named $C_{\mathrm{Hyp},i}$ and $C_{\mathrm{BS},i}$. We also compute the (unweighted) median and standard deviation of the difference of the absolute errors. Positive values for the weighted mean and the median are obtained when a GH model produces a smaller pricing error. The hitting rate gives the percentage of trading volume for which the Black-Scholes price is closer to the market price.

The weighted means (2.5) and the median of the difference of the absolute errors are positive for all volatility estimators and subclasses of GH distributions, hence we always observe an improvement in comparison to the Black-Scholes model. This is confirmed by the values of the hitting rate, which are about 31% for the historical volatility estimator and about 42% for the implicit volatility estimators. The best results under all GH models with respect to the hitting rate are obtained for the symmetric centered hyperbolic model whereas the largest improvements to the Black-Scholes benchmark in terms of the difference of absolute errors are observed for the GH model with arbitrary $\lambda$.

Table 2.20: *Comparison of the pricing errors for Bayer calls. "SC" marks the symmetric centered versions of the models.*

| Estimator | Models | Difference of the abs. errors | | | Hitting Rate BS model |
|---|---|---|---|---|---|
| | | w. mean | median | st. dev. | |
| Hist30 | BS / Hyperbolic | 0.0058 | 0.0065 | 0.0189 | 31.2419 |
| Hist30 | BS / Hyperbolic SC | 0.0112 | 0.0137 | 0.0271 | 29.5022 |
| Hist30 | BS / NIG | 0.0131 | 0.0163 | 0.0353 | 30.0227 |
| Hist30 | BS / NIG SC | 0.0088 | 0.0098 | 0.0281 | 31.4254 |
| Hist30 | BS / GH | 0.021 | 0.0264 | 0.0594 | 32.258 |
| Hist30 | BS / GH SC | 0.016 | 0.0181 | 0.0483 | 31.645 |
| Imp.median20 | BS / Hyperbolic. | 0.0018 | 0.0004 | 0.0547 | 42.0699 |
| Imp.median20 | BS / Hyperbolic SC | 0.0024 | 0.0006 | 0.0542 | 41.4553 |
| Imp.median20 | BS / NIG | 0.0027 | 0.0008 | 0.0783 | 41.8004 |
| Imp.median20 | BS / NIG SC | 0.0024 | 0.0006 | 0.0788 | 42.0026 |
| Imp.median20 | BS / GH | 0.0032 | 0.0008 | 0.1216 | 43.3419 |
| Imp.median20 | BS / GH SC | 0.0032 | 0.0008 | 0.1217 | 43.4586 |
| Imp.median30 | BS / Hyperbolic | 0.0014 | 0.0005 | 0.046 | 42.4558 |
| Imp.median30 | BS / Hyperbolic SC | 0.0021 | 0.0007 | 0.0477 | 41.9724 |
| Imp.median30 | BS / NIG IL | 0.0024 | 0.0009 | 0.0652 | 42.5437 |
| Imp.median30 | BS / NIG SC | 0.0019 | 0.0007 | 0.0645 | 42.4173 |
| Imp.median30 | BS / GH | 0.0023 | 0.0013 | 0.0987 | 43.2795 |
| Imp.median30 | BS / GH SC | 0.0023 | 0.0012 | 0.0987 | 43.2614 |
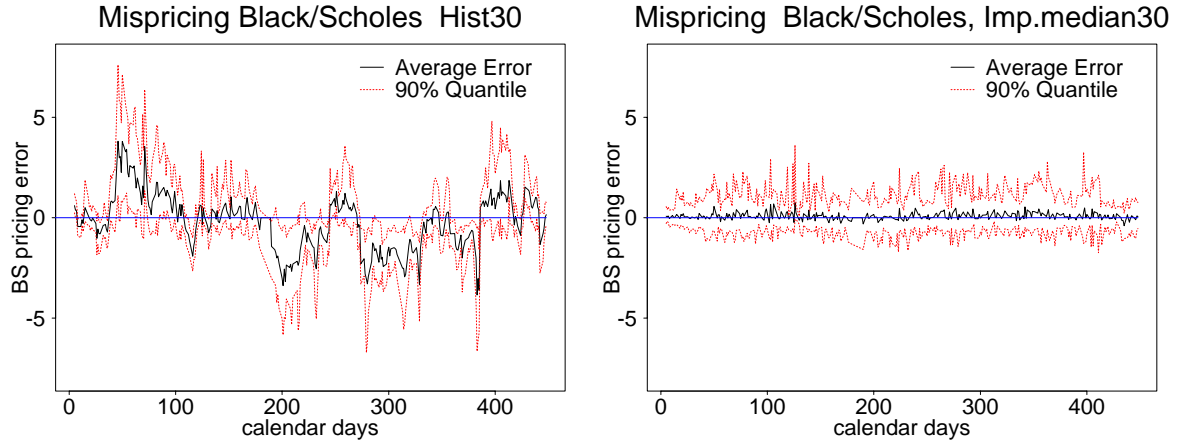
Figure 2.21: *Pricing performance of Thyssen calls as a time series, June 1, 1993 to August 18, 1994. The average and quantiles are computed on a daily basis.*

## When do we Observe Mispricing?

In the first part of this section we analysed pricing patterns with respect to time-to-maturity and stockprice-strike ratio. For instance we observed small pricing errors for at-the-money options with short maturities. At first sight, the average daily pricing errors as a time series in Figure 2.21 reveals that Black-Scholes prices computed with intraday implicit volatilites are less erroneous than prices based on historical volatility estimates. The results for the prices correspond to the volatility patterns in Figure 2.16.

To give a more explicit answer to the question in the title of this section, we look at the correlations of the average daily pricing errors and their standard deviations with other observable parameters, respectively. Below we describe only those results which hold for both, Bayer and Thyssen options and both volatility estimators Imp.median20 and Hist30. Although the volatility estimators used for the calculation of prices cause errors with distinct patterns, we obtain similar results concerning the correlations with other observables.

First, we analyze the correlation of the errors with three volatility estimates: the average daily implicit volatility (denoted by ImpVol), the historical volatility estimator Hist30, and the volatility index VDAX of the Deutsche Terminbörse. The index VDAX is constructed as the implicit volatility of at-the-money options on the German stock index DAX with 45 days to expiration. Obviously the spread of the pricing errors SdErr has a positive correlation with all three volatility series.

Looking at the correlation of the daily trading volume and the standard deviations of the pricing errors, reveals that at those days with an increased trading it is also more difficult to compute prices which are close to the market. Moreover, the spread of the pricing errors increases if the stockprice (Kassakurs) increases.

Table 2.22: *Correlation of the average daily pricing errors AvErr (resp. standard deviations SdErr of daily pricing errors) of Black-Scholes prices and other parameters. The p-values are computed using Pearson's test with a 2-sided alternative. We applied the volatility estimators Hist30 and Imp.median20. ImpVol is the average value (weighted by the trading volume) of the Black-Scholes implicit volatilities during one trading day and Kassakurs the value of the stock at 12 o'clock.*

| | | Bayer calls | | Thyssen calls | |
|---|---|---|---|---|---|
| | | Correlation | p-value | Correlation | p-value |
| **Hist30** | | | | | |
| AvErr | ImpVol | −0.13 | 0.0025 | −0.108 | 0.0149 |
| SdErr | ImpVol | 0.124 | 0.0040 | 0.19 | 0 |
| AvErr | Hist30 | 0.582 | 0 | 0.684 | 0 |
| SdErr | Hist30 | 0.379 | 0 | 0.502 | 0 |
| AvErr | VDAX | 0.357 | 0 | −0.126 | 0.0029 |
| SdErr | VDAX | 0.452 | 0 | 0.353 | 0 |
| AvErr | Kassakurs | 0.206 | 0 | 0.02 | 0.5664 |
| SdErr | Kassakurs | 0.626 | 0 | 0.615 | 0 |
| AvErr | Volume | 0.062 | 0.0846 | 0.019 | 0.6020 |
| SdErr | Volume | 0.34 | 0 | 0.412 | 0 |
| **Imp.median20** | | | | | |
| AvErr | ImpVol | 0.076 | 0.0785 | −0.2 | 0 |
| SdErr | ImpVol | 0.242 | 0 | 0.285 | 0 |
| AvErr | Hist30 | −0.087 | 0.0468 | −0.013 | 0.7710 |
| SdErr | Hist30 | 0.035 | 0.4219 | 0.148 | 0.0011 |
| AvErr | VDAX | 0.024 | 0.5771 | −0.009 | 0.8327 |
| SdErr | VDAX | 0.221 | 0 | 0.267 | 0 |
| AvErr | Kassakurs | 0.251 | 0 | 0.069 | 0.0479 |
| SdErr | Kassakurs | 0.56 | 0 | 0.653 | 0 |
| AvErr | Volume | 0.198 | 0 | 0.009 | 0.8050 |
| SdErr | Volume | 0.38 | 0 | 0.518 | 0 |

## 2.6 Statistical Martingale Measures

In the Black-Scholes model the price measure depends only on the volatility of the geometric Brownian motion. Thus, in this case the estimation of the risk-neutral density implicit in the secondary market corresponds to the computation of implicit volatilities (cf. Section 2.4). See Christensen and Kiefer (1999) for a simulation approach to martingale measures in complete continuous time diffusion models. Here we assume that log-prices follow a GH Lévy motion also under the martingale measure.

In the GH model the risk-neutral density is characterized by more than one parameter. Consequently, if we assume that the risk-neutral distribution is generalized hyperbolic, it is not possible to estimate the density from a single quote observed at the secondary market. Recall, that we fixed the shape of the distribution in the preceding sections whereas we are interested in the shape of the risk-neutral distribution in this section. However, in the following way it is possible to estimate the risk-neutral GH density from a set of observed option prices: Following Keller (1997, pp. 97–105), Eberlein, Keller, and Prause (1998) in models where the log-price follows a Lévy process, the risk-neutral distribution has to be chosen such that $r = \log M(1)$ is satisfied, where $M$ is the moment-generating function of $\mathcal{L}(X_1)$. Recall, that a homogeneous Lévy process $(X_t)_{t \geq 0}$ is characterized by the distribution of $X_1$. Therefore, it is necessary to find the optimal GH density in the space $\Theta$ of parameters which yield a martingale measure

$$
\Theta = \Big\{ (\lambda, \alpha, \beta, \delta, \mu) \,\Big|\, \lambda \in \mathbb{R}, \quad 0 < |\beta| < \alpha, \quad \delta > 0,
$$
$$
\mu = r - \log \frac{K_\lambda(\delta \sqrt{\alpha^2 - (\beta + 1)^2})}{K_\lambda(\delta \sqrt{\alpha^2 - \beta^2})} + \frac{\lambda}{2} \log \frac{\alpha^2 - (\beta + 1)^2}{\alpha^2 - \beta^2} \Big\}.
$$

Note, that statistical martingale measures (SMM) are not necessarily equivalent to objective measures estimated from time series data. Moreover, since option traders could assume a price process different to the empirically observed process, the equivalence of the measure to the empirically observed probability measure is for the computation of statistical martingale measures not relevant.

In particular, Keller (1997) proposed to minimize (squared) pricing errors, i.e. to find the statistical martingale measure as

$$
\mathrm{argmin}_{(\lambda, \alpha, \beta, \delta, \mu) \in \Theta} \sum_i \left| \widehat{C}_i - C_{\mathrm{model}}(\vartheta) \right|^h, \tag{2.6}
$$

where $h = 1, 2$. We call the obtained probability measure a *pricing error minimizing* SMM. See also Özkan (1997) for latter SMM estimates in the hyperbolic model.

The minimization of (2.6) yields an optimal SMM with respect to pricing performance. Another aspect in the testing of option pricing models, equally important to pricing, is the smile of implicit volatilities. This leads to the question which SMM explains the smile effect, i.e. for which risk-neutral shape parameters do we get a flat implicit volatility curve with respect to moneyness and time to maturity? To find an answer, we minimize

$$
\mathrm{argmin}_{(\lambda, \chi, \psi) \in \Theta} \sum_i \left| \sigma_{\mathrm{implicit}}(\widehat{C}_i, \lambda, \chi, \psi) - \overline{\sigma_{\mathrm{implicit}}} \right|^h, \tag{2.7}
$$

where $h = 1, 2$ and $\overline{\sigma_{\mathrm{implicit}}}$ denotes the average implicit volatility with respect to the Esscher transformed historical parameter estimates and $\Theta$ is given in the 3rd parametrization (1.3).

The implicit volatility $\sigma_{\text{implicit}}(\widehat{C}_i, \lambda, \chi, \psi)$ of each quote is computed as in Section 2.4. Note, that we minimize (2.7) only w.r.t. scale- and location-invariant parameters. For practical purposes the value of $\overline{\sigma_{\text{implicit}}}$ has only a minor influence on the obtained scale- and location-invariant parameters: We have replaced $\overline{\sigma_{\text{implicit}}}$ by other values, e.g. the median of the implicit volatility and a volatility estimate from time series data; the results do not change. Note, that the scale of these risk-neutral distributions plays no role for the calculation of the distance (2.7), since we always rescale the distribution to compute implicit volatilities. We use the rescaling mechanism described in Section 2.1. Therefore, these risk-neutral distributions are rescaled to the variance of the time series estimates. This SMM, which we call *smile minimizing*, describes the shape, i.e. the skewness and kurtosis of the return distribution assumed by traders in the secondary market.

One could also use shape parameters of smile minimizing SMMs to construct a volatility estimator based on GH implicit volatilities. Usually the smile is removed by using only options at-the-money with a given time to maturity to construct a volatility estimator (see for instance Deutsche Börse (1997) for the removal of the smile effect in the construction of the VDAX). An estimator based on the smile minimizing martingale measure minimizes the noise in Black-Scholes implicit volatility estimators coming from the smile effect (see Eberlein, Keller, and Prause (1998) for a similar approach based on the Esscher martingale measure).

## Computational Aspects

We apply the Powell algorithm implemented by Özkan (1997) to estimate the SMMs by minimizing (2.6) or (2.7). In the case of the statistical martingale measure computed by minimizing pricing errors, a further refinement of the minimization algorithm allows a much faster computation: the first step in the refined minimization procedure is to fix the shape of the GH distribution and to change only the scale parameter. Since the choice of the volatility is the crucial point in the computation of prices which are close to observed market prices, the speed of the minimization procedure increases markedly.

We estimate the statistical martingale measure only for the subclass of NIG distributions. Since NIG distributions are closed under convolution we have an explicit risk-neutral density for all times to maturity. Hence, we avoid possible numerical errors from the Fast Fourier Transform for extreme parameter values of the GH distribution.

## Statistical Results

Figure 2.23 shows the densities of daily log-returns of Bayer stocks estimated in different ways. The kernel-density estimate and likelihood estimate are obtained from 6 years of observations. The parameters of all the GH estimates are given in Table 2.27. Obviously, the likelihood estimate is the GH distribution closest to the empirical data.

The densities of the pricing error minimizing SMM (for the $L^1$ and $L^2$ distance) differ markedly from the estimated densities based on time series data. Both densities are skewed, in particular they put more weight in the right tail than in the left. Recall, that skewed risk-neutral GH densities yield a larger difference to Black-Scholes prices also for longer maturities (see Figure 2.6). Taken together, the pricing error minimizing SMMs have less kurtosis than the likelihood estimate. This is also highlightened by their position in the shape triangle plotted in Figure 2.24. Both positions are not close to the usually estimated values of $\xi$
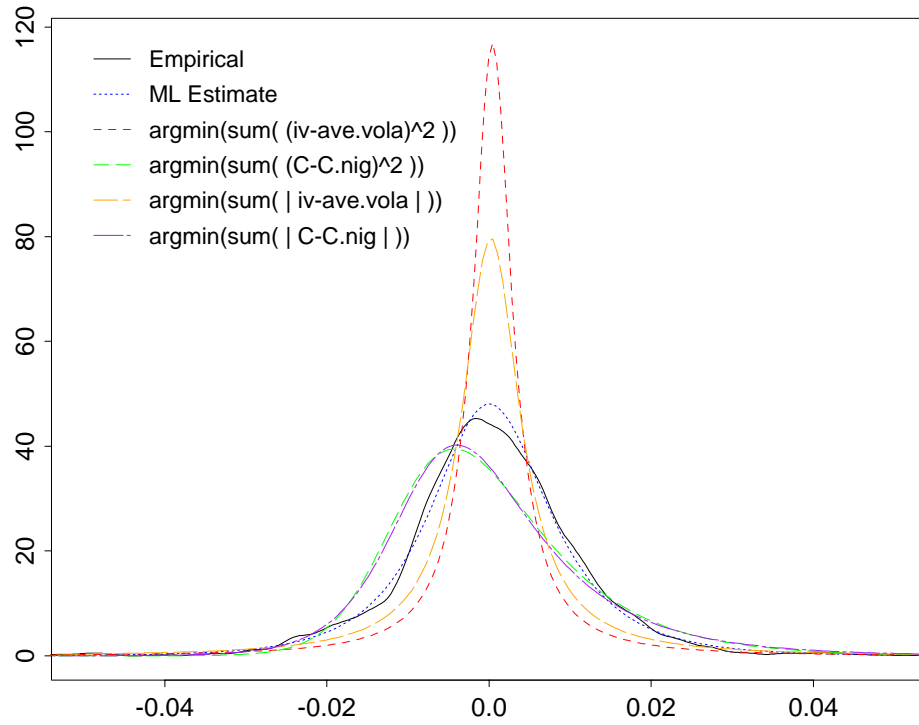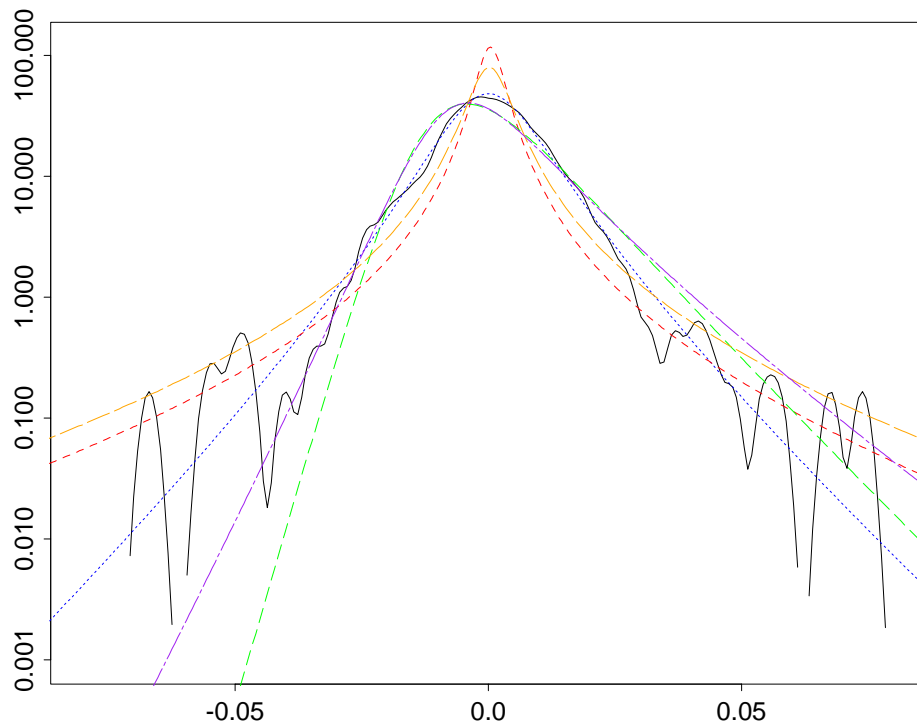
Figure 2.23: *Comparison of empirical and GH densities obtained as likelihood estimates and as SMM estimates: Bayer stock and call options, January 1994, 1240 observations.*

and $\chi$ for financial assets. See Barndorff-Nielsen and Prause (1999) for positions of NIG distributions estimated for various financial assets in the shape triangle.

A distinct picture is obtained for the smile minimizing SMM. Densities estimated in this way show a substantial increase in the kurtosis. Note, that the smile minimizing risk-neutral densities in Figure 2.23 are rescaled, such that their variance equals the variance of the likelihood estimate. The high kurtosis is also visible in the NIG shape triangle in Figure 2.24; the position of the smile minimizing SMM is close to the top of the triangle, above almost all $\xi$-values estimated for historical data. However, in contrast to the pricing error minimizing SMMs, the smile minimizing risk-neutral distributions are approximately centered.

In Figure 2.25 we have plotted the IG distributions obtained by the mixing representation (1.6). A natural way to construct stochastic volatility models based on the GH model is to look at these mixing IG distributions (see Chapter 3, Barndorff-Nielsen and Shephard 1999, Nicolato and Prause 1999). Moreover, the independent IG distributions may be replaced by a more appropriate process. Figure 2.25 shows that volatility processes with IG marginals obtained from the various SMM measures should differ substantially from the mixing IG distributions inherent in likelihood estimates.

In Figure 2.26 we have plotted the implicit volatilities obtained for the smile minimizing martingale measure. Obviously, by choosing appropriate shape parameters, the smile is markedly reduced, although it is not totally removed (e.g. compare Figure 2.26 with Figures 2.13).

Another way to show this is to fit a linear model as in Section 2.4. We estimate the coefficients in the same linear model (2.3) for Bayer calls from July 1992 to August 1994. While we observe a coefficient $b_2 = 51.77$ in the Black-Scholes model and $b_2 = 46.13$ in the symmetric and centered GH model, we obtain $b_2 = 34.32$ for the implicit volatility in the smile minimizing martingale measure.

## Pricing Performance

Finally, we use the SMM estimates to compute prices and compare them with observed prices. To compute a particular price, we fix the shape parameters of the GH distributions and rescale the distribution with short-term volatility estimates, i.e. we follow the same procedure as in Section 2.5. The SMM estimates for the shape are obtained from Bayer call options in January 1994. We use these SMM estimates to compute prices for the whole two year period in the same way as in Section 2.5. The results are given in Table 2.28.

First we look at the pricing error minimizing martingale measures: with one exception (Hist30, $L^1$) they have a mean error which is smaller in absolute terms but a higher standard deviation than the Black-Scholes prices.

On the contrary, the pricing performance of the smile-minimizing martingale measure is worse than the performance of the corresponding Black-Scholes approach. Both, mean error and standard deviation increase in this NIG SMM approach. Of course, the poor results for the smile-minimizing SMMs are not surpricing.

Taken together, SMMs do explain aspects of pricing behaviour and implicit volatility patterns, but both approaches to find a martingale measure do not coincide at all.
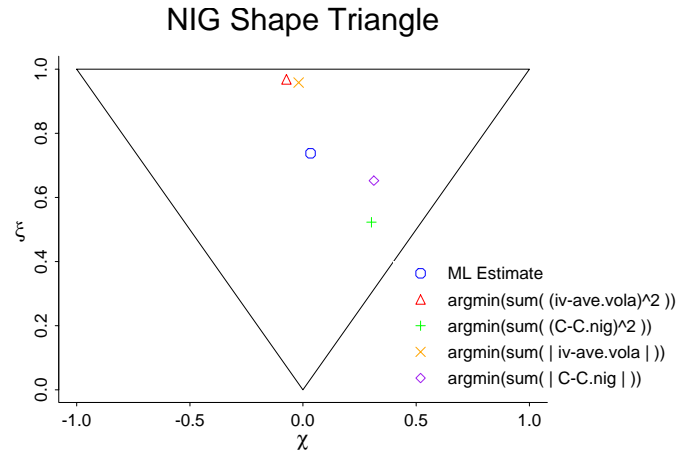
Figure 2.24: *Comparison of historical estimates and statistical martingale measures. Bayer calls, January 1994, 1240 observations.*
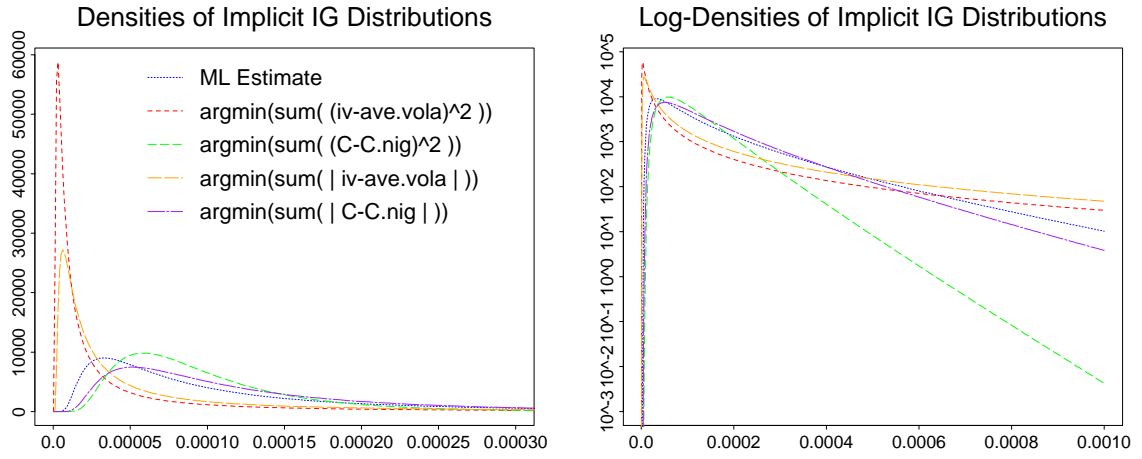


Figure 2.25: *Implicit GIG distributions. Bayer calls, January 1994, 1240 observations.*
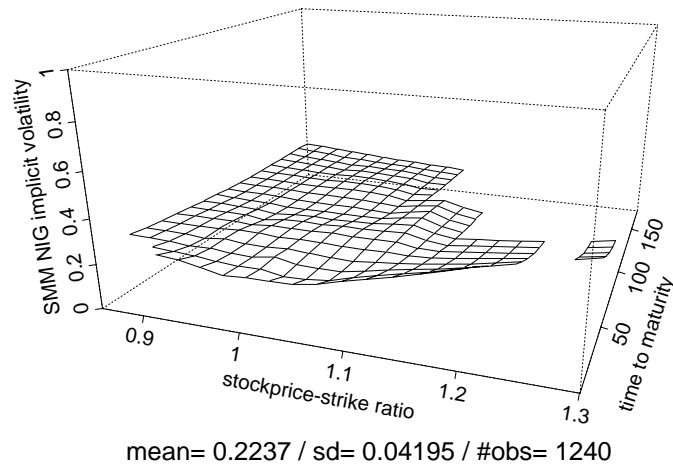


mean= 0.2237 / sd= 0.04195 / #obs= 1240

Figure 2.26: *NIG implicit volatilities for the statistical martingale measure calculated as* $\mathrm{argmin}\left(\sum(iv\text{-}ave.vola)^2\right)$. *Bayer calls, January 1994.*

64

Table 2.27: *Historical and SMM normal inverse Gaussian densities. Bayer stock 1988-94 and Bayer call options, January 1994.*

| Method | $\alpha$ | $\beta$ | $\delta$ | $\mu$ | $\xi$ | $\chi$ | Vola |
|---|---|---|---|---|---|---|---|
| ML estimate | 81.6 | 3.69 | 0.0103 | $-0.00012$ | 0.74 | 0.033 | 0.178 |
| *Smile minimizing SMM* | | | | | | | |
| $\mathrm{argmin}\big(\sum(\text{iv-ave.vola})^2\big)$ | 22.9 | $-1.73$ | 0.0029 | 0.00042 | 0.97 | $-0.07$ | 0.178 |
| $\mathrm{argmin}\big(\sum|\text{iv-ave.vola}|\big)$ | 26.4 | $-0.53$ | 0.0034 | 0.00022 | 0.96 | $-0.02$ | 0.178 |
| *Pricing error minimizing SMM* | | | | | | | |
| $\mathrm{argmin}\big(\sum(\widehat{C}-\mathrm{C}_{\mathrm{NIG}})^2\big)$ | 200.2 | 115.96 | 0.0163 | $-0.01144$ | 0.52 | 0.30 | 0.194 |
| $\mathrm{argmin}\big(\sum|\widehat{C}-\mathrm{C}_{\mathrm{NIG}}|\big)$ | 114.6 | 55.09 | 0.0134 | $-0.00722$ | 0.65 | 0.31 | 0.208 |

Table 2.28: *Pricing performance of the statistical martingale measure calculated for the Bayer calls January 1994 (compare with Table 2.27). The prices are computed for Bayer calls, July 1, 1992 – August 19, 1994, 21021 observations.*

| **Bayer calls** | | $C_{\mathrm{model}} - C_{\mathrm{quoted}}$ | |
|---|---|---|---|
| model | volatility estimator | mean | st.dev. |
| *Pricing error minimizing SMM, L1, Bayer calls Jan 1994* | | | |
| NIG | Imp.median20 | 0.06 | 2.062 |
| BS | Imp.median20 | 0.279 | 1.475 |
| NIG | Imp.median30 | 0.063 | 1.979 |
| BS | Imp.median30 | 0.276 | 1.399 |
| NIG | Hist30 | 1.335 | 2.114 |
| BS | Hist30 | 0.798 | 2.389 |
| *Pricing error minimizing SMM, L2, Bayer calls Jan 1994* | | | |
| NIG | Hist30 | 0.784 | 2.402 |
| Black-Scholes | Hist30 | 0.798 | 2.389 |
| NIG | Imp.median20 | $-0.222$ | 2.64 |
| BS | Imp.median20 | 0.279 | 1.47 |
| NIG | Imp.median30 | $-0.224$ | 2.57 |
| BS | Imp.median30 | 0.276 | 1.4 |
| *Smile minimizing SMM, L2, Bayer calls Jan 1994* | | | |
| NIG | Imp.median20 | 0.309 | 1.46 |
| Black-Scholes | Imp.median20 | 0.279 | 1.475 |
| NIG | Imp.median30 | 0.311 | 1.444 |
| Black-Scholes | Imp.median30 | 0.276 | 1.399 |
| NIG | Hist30 | 0.564 | 2.328 |
| Black-Scholes | Hist30 | 0.798 | 2.389 |

## 2.7 Alternative Testing Approaches

Lo (1986) developed a methodology which allows for a test of an option pricing model in a classical statistical sense. He proposed to find confidence intervals for the estimates of the parameters in the model of interest. In particular the volatility is the most important parameter. The variance may be estimated by the canonical ML-estimator, which is a consistent and uniformly asymptotically normal distributed (CUAN) estimator (Pfanzagl 1994). Since option prices are a monotone function with respect to volatility, this results in confidence bounds for the option price (for a specified model). The test, build on large-sample theory, has the following deficiency: Since in reality frequent changes in the volatility of the underlying do occur, it is not appropriate to use arbitrarily large data sets for the estimation of volatility parameters. Another deficiency comes from the fact that quoted option prices reflect the behaviour of stock prices expected by the traders for the period from trade to expiration. This is not the period of time used to estimate historical parameters. The trader's forecasts may be biased in an unknown way, e.g. in expectation of a crash which will never occur. Hence, a rigorous statistical test as described above is difficult if not impossible. Moreover, the differences in the underlying and the risk-neutral density were used to test for crash fears in the secondary market (Bates 1991; Bates 1997; Christensen and Prabhala 1998). Therefore, this asymptotic approach seems not to be adequate to test the performance of option pricing models. See Figure 2.29 for an illustration of the problems in the construction of tests for option pricing models.
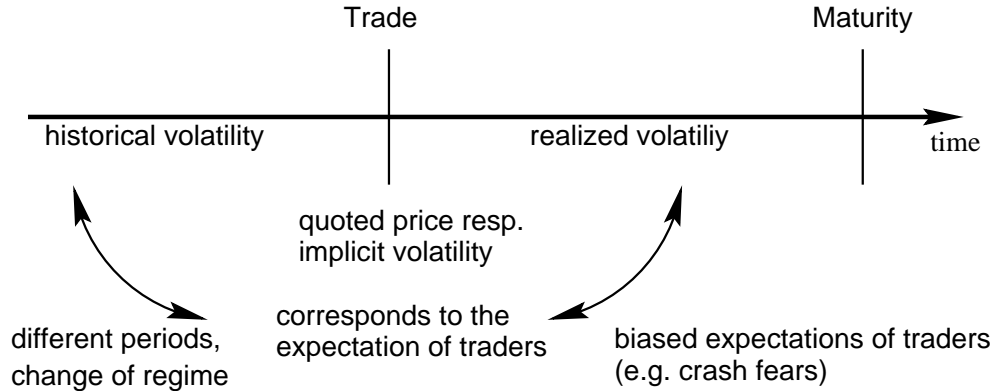


Figure 2.29: *Fundamental problems in the evaluation of option pricing models.*

Bakshi, Cao, and Chen (1998) observed that in contrast to the monotonicity of call prices with respect to stock prices, which is assumed in all (reasonable) models with one source of randomness, stock prices and call prices often move in opposite directions. The same result holds for the German market. We zoomed down in single option series and calculated the increments $\Delta C_i$ and $\Delta S_i$ of stocks and call options from quote to quote. Overnight-increments were removed to avoid decreasing call prices due to the shorter time to maturity. For roughly 10% of the quoted increments the product $\Delta C_i \Delta S_i$ is negativ (type-I violations in Bakshi, Cao, and Chen (1998)). From Table 2.30 we also see that the (absolute) magnitude of the call and stock price violating the monotonicity of the increments is clearly above tick-size. These violations hint at possible difficulties for the practical implementation of hedging strategies.

There are different possible explanations for this effect. Bakshi, Cao, and Chen (1998) concluded from their empirical results that a stochastic volatility model may explain this

Table 2.30: *Do call prices and stock prices move in different directions? Percentage and magnitude of violations of the monotonicity in Daimler Benz call series. The last four columns refer only to those quotes violating the monotonicity condition.*

| strike | maturity | | quotes | errors | | mean abs. incr. | | median abs. incr. | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | # | % | call | stock | call | stock |
| 550 | 1 | 93 | 641 | 64 | 9.984 | 0.4125 | 1.0422 | 0.2 | 1 |
| 700 | 12 | 92 | 254 | 26 | 10.236 | 0.4885 | 1.2308 | 0.5 | 1 |
| 800 | 9 | 92 | 138 | 14 | 10.145 | 0.2071 | 1.1286 | 0.2 | 0.5 |
| 800 | 12 | 92 | 142 | 16 | 11.268 | 0.425 | 0.9875 | 0.45 | 0.65 |
| 800 | 10 | 93 | 511 | 70 | 13.699 | 0.2743 | 1.1229 | 0.2 | 1 |
| 900 | 12 | 94 | 133 | 19 | 14.286 | 0.5579 | 1.1474 | 0.4 | 1 |

effect; but, they also mention that the Hull and White (1987) model of stochastic volatility only explains 60% of the observed violations of the monotonicity. However, if these violations could be explained by stochastic volatility models, the volatility of the average volatility during the expiration period of the call must be rather high. A more appropriate explanation of these violations is the influence of trading at the secondary market. Prices of derivative assets are not only determined by the behaviour of the underlying but also by supply and demand.

## 2.8   Conclusion

This chapter contains a thorough empirical study of the implicit volatility patterns and the pricing behaviour of the proposed generalized hyperbolic model and the Black-Scholes model. The conclusion from the empirical results is that GH models improve the Black-Scholes model, but they do not explain all deficiencies of the Black-Scholes model. To get this complete picture of the advantages and disadvantages it is necessary to perform more than a single test.

It is worth to mention that with respect to option pricing new models lead to intricate problems. On the one hand it is impossible to take care of all kinds of "stylized features" within one particular model. On the other hand the statistical and numerical problems which arise are often substantial.

The problems which are not solved by any of the described models are on the one hand a realistic modelling of the depencency structure and on the other hand the trading effects from the secondary market. Stochastic volatility and trading effects are possible reasons for the smile and smirks of implicit volatilities and the violation of the monotonicity condition of stock and option movements.

# Chapter 3

# Stochastic Volatility Models of the Ornstein-Uhlenbeck type

Following the results of the last chapter we have to refine the GH Lévy model by an appropriate modelling of the volatility structure. This is done in a natural way by assuming that the latent volatility is a process of the Ornstein-Uhlenbeck type. This model, proposed by Barndorff-Nielsen (1998), Barndorff-Nielsen and Shephard (1999), is analytically as well as numerically tractable.[1] Parts of this chapter are a joint work with Elisa Nicolato; the results are published in Nicolato and Prause (1999).

## 3.1  Empirical Motivation

In the sequel we discuss volatility processes with GIG, and more precisely, IG and Gamma marginals. Wirth (1998) has shown that GIG distributions are an appropriate model for the marginals of the German volatility index VDAX and also for the marginals of historical volatility estimates. As a motivation for the introduction of volatility processes with IG marginals, we compare in Figure 3.1 daily averages of Black-Scholes implied volatilities of Bayer stock options and mixing IG distribution of the NIG estimate from Bayer stock returns. The mixing IG distributions are implicitly given for known NIG distributions by the variance-mean mixture (1.6). Black-Scholes implied volatilities are often markedly different to historical volatility estimates. This is in particular true for short time periods. Therefore, it is no surprise that the empirical distribution of Black-Scholes implicit volatilities given in Figure 3.1 differs substantially from the mixing IG distribution. However, one could estimate an IG distribution which fits Black-Scholes implied volatilities precisely. This IG estimate is also shown in Figure 3.1. For the estimation we use an algorithm implemented by Wirth (1998).

In Figure 3.2 we show GIG distributions corresponding to one- and three-dimensional estimates of hyperbolic distributions. Note, that the scale of mixing GIG distributions differ

---

[1]We have to add some remarks concerning the notation. The parameters used in this chapter are consistent in this dissertation and with most papers concerning the GH distributions. However, they differ from the parametrizations used in Barndorff-Nielsen (1998), Barndorff-Nielsen and Shephard (1998, 1999), Nicolato and Prause (1999) in the following way: In this study $\lambda$ denotes the subclass and $\tau$ describes the autocorrelation (instead of $\bar{\lambda}$ and $\lambda$). We changed the parametrizations of the GIG distributions in this chapter (see equation (3.15)).
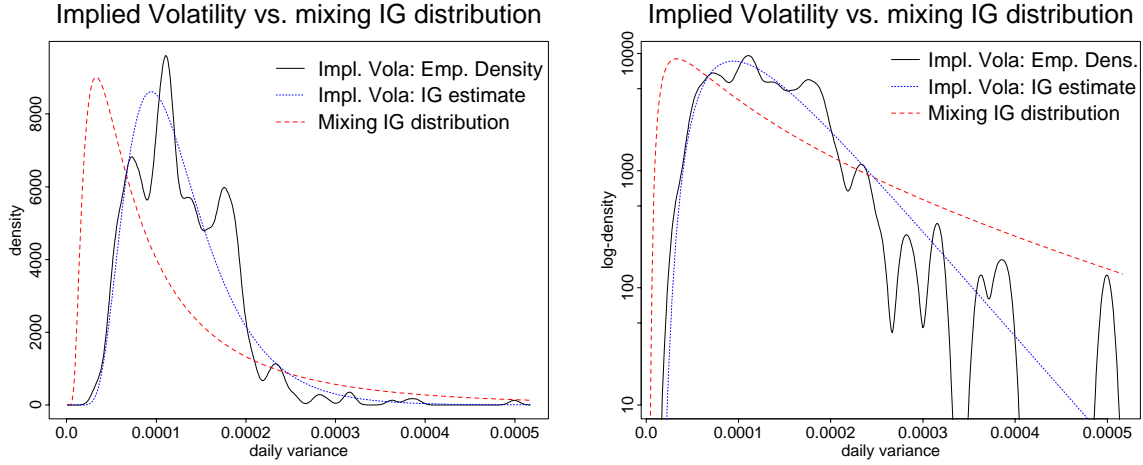
Figure 3.1: *Comparison of the empirical density and the IG likelihood estimate of Black-Scholes implied volatilities with mixing IG distributions (Bayer stocks 1988–94 and stock options 1992–94).*
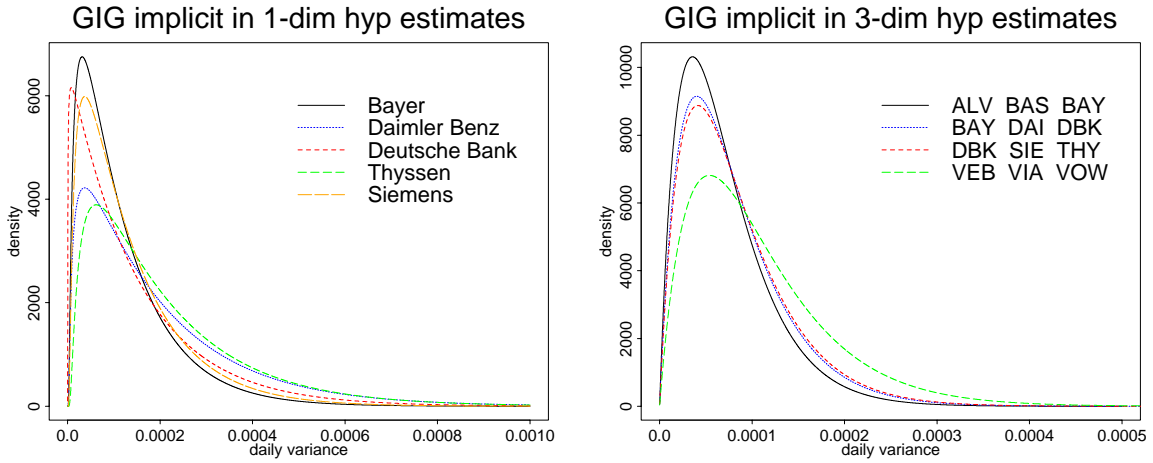


Figure 3.2: *Generalized inverse Gaussian distributions implicit in estimates of univariate and multivariate hyperbolic distributions.*

from the scale of the typically examined volatility (annualized standard deviations) since we look at daily variances.

In Section 2.2 we have already mentioned that a significant autocorrelation is usually observed in the squared or absolute returns of financial time series. The model we are investigating in this chapter has the (quasi-)long-range dependence typically observed in volatilites of financial time series (see Barndorff-Nielsen (1998) for the details).

## 3.2 Ornstein-Uhlenbeck type Processes

First, we give a brief introduction to Ornstein-Uhlenbeck type processes which were introduced by Barndorff-Nielsen (1998) as a model in Finance. Further references to Ornstein-Uhlenbeck type processes are Wolfe (1982), Sato and Yamazato (1982), Jurek and Vervaat (1983), Sato, Watanabe, and Yamazato (1994). These processes can be constructed on the basis of the following class of distributions.

**Definition 3.1 (Self-Decomposability).** *A probability measure $\mu$ (or equivalently its characteristic function $\varphi$) is defined as* self-decomposable *(or to belong to Lévy's class L), if for each $\tau > 0$ there exists a probability measure $\nu_\tau$ such that*

$$\varphi(\zeta) = \varphi(e^{-\tau}\zeta)\varphi_\tau(\zeta), \tag{3.2}$$

*where $\varphi$ and $\varphi_\tau$ denote the characteristic functions of $\mu$ and $\nu_\tau$ respectively.*

Note, that self-decomposable distributions are infinitely divisible. The following characterization clarifies the structure of self-decomposable distributions.

**Theorem 3.3 (Characterization of Self-Decomposable Distributions).** *A    random variable $Y$ has a self-decomposable distribution if and only if*

$$Y = \int_0^\infty e^{-t}dZ_t, \tag{3.4}$$

*where $(Z_t)_{t\geq0}$ is a Lévy process. The relation of the Lévy measures $V$ and $W$ of $Y$ and $Z_1$ is then given by*

$$V(dy) = \int_0^\infty W(e^t dy)dt. \tag{3.5}$$

*If there exist continuous Lebesgue densities $v$ and $w$ for $V$ and $W$ respectively, they satisfy*

$$v(y) = y^{-1}W([x, \infty)), \tag{3.6}$$

$$w(y) = -v(y) - yv'(y), \quad y > 0. \tag{3.7}$$

The theorem was proved by Jurek and Vervaat (1983), see also Barndorff-Nielsen (1998, Theorem 2.2).

**Definition 3.8 (Ornstein-Uhlenbeck type Processes).** *A stochastic process $(Y_t)_{t\geq0}$ is defined as a process of the* Ornstein-Uhlenbeck type *if it satisfies the stochastic differential equation*

$$dY_t = -\tau Y_t dt + dZ_t, \tag{3.9}$$

*where $\tau > 0$ and $(Z_t)_{t\geq0}$ is a Lévy process, to which we refer to as the* background driving Lévy process *(BDLP).*

71

We take a càdlàg version of the process $(Z_t)_{t \geq 0}$ and assume $\tau > 0$. The stochastic differential equation (3.9) is then solved by

$$Y_t = e^{-\tau t} Y_0 + \underbrace{\int_0^t e^{-\tau(t-s)} dZ_s}_{=: \, u_t}. \tag{3.10}$$

where $Y_0$ and the BDLP are independent. With this definition of $u_t$ we may write $Y_t = \exp(-\tau t) Y_0 + u_t$. If the solution (3.10) is stationary and square integrable and if $\mathrm{E}[x_0] = \mathrm{E}[Z_1] = 0$, then the correlation function of $(Y_t)_{t \geq 0}$ is of the form $\mathrm{acf}(u) = \exp(-\tau u)$.

The following theorem yields the existence of a stationary solution $(Y_t)_{t \geq 0}$, with a marginal distribution given by the self-decomposable distribution.

**Theorem 3.11 (Existence of Ornstein-Uhlenbeck type Processes).** *Let $c(\zeta)$ be a differentiable and self-decomposable characteristic function and let $\kappa(\zeta) = \log c(\zeta)$. Suppose, that $\zeta \kappa'(\zeta)$ is continuous at 0 and let $\varphi^\tau(\zeta) = \tau \zeta \kappa'(\zeta)$ for a $\tau > 0$. Then $\exp\big(\varphi^\tau(\zeta)\big)$ is an infinitely divisible characteristic function.*

*Furthermore, letting $(Z_t^\tau)_{t \geq 0}$ be the Lévy process for which $Z_1^\tau$ has the characteristic function $\exp\big(\varphi^\tau(\zeta)\big)$ and defining the process $(Y_t)_{t \geq 0}$ by $dY_t = -\tau Y_t dt + dZ_t^\tau$, we assert that a stationary version of $Y$ exists, with one-dimensional marginal distribution given by the characteristic function $c(\zeta)$.*

The theorem was proved by Barndorff-Nielsen (1998, Theorem 3.2), see also Barndorff-Nielsen, Jensen, and Sørensen (1995, Section 3). Since $(Z_t^\tau)_{t \geq 0}$ is a Lévy process, $\varphi^\tau(\zeta) = \tau \varphi^1(\zeta)$ holds and we have that $(Z_t^\tau)_{t \geq 0}$ is identical in law to $(Z_{\tau t}^1)_{t \geq 0}$. Using the abbreviation $Z_t := Z_t^1$ we have

$$(Z_t^\tau)_{t \geq 0} \overset{\mathcal{L}}{=} (Z_{\tau t})_{t \geq 0}, \tag{3.12}$$

where $\overset{\mathcal{L}}{=}$ means "equal in law". This implies that

$$Y_t \overset{\mathcal{L}}{=} e^{-\tau t} Y_0 + e^{-\tau t} \int_0^{\tau t} e^s dZ_s \tag{3.13}$$

with the corresponding differential equation

$$dY_t = -\tau Y_t dt + dZ_{\tau t}. \tag{3.14}$$

### Inverse Gaussian Ornstein-Uhlenbeck type Processes

The $\mathrm{IG}(\delta, \gamma)$ distribution is given by the Lebesgue density[2]

$$\frac{\delta}{\sqrt{2\pi}} \, \exp(\delta \gamma) \, x^{-3/2} \, \exp\left( -\frac{1}{2}\big(\delta^2 x^{-1} + \gamma^2 x\big) \right), \quad x > 0. \tag{3.15}$$

GIG and in particular IG distributions are self-decomposable (Halgreen 1979). Therefore, we can construct a GIG Ornstein-Uhlenbeck type process $(Y_t)_{t \geq 0}$, i.e. a stationary process such that $Y_t \sim \mathrm{GIG}$ for all $t \geq 0$. To conserve analytical and numerical tractability, we restrict

---

[2]This notation is related by $\chi = \delta^2$ and $\psi = \gamma^2$ to the notation in Chapter 1.

ourselves to IG and Gamma distributions (the processes are denoted as IGOU or Gamma-OU, for short). However, most of the interesting results hold also for GIG Ornstein-Uhlenbeck type processes. The Lévy density of IG distributions is

$$v(x) = \frac{1}{\sqrt{2\pi}} \delta \, x^{-3/2} \, \exp(-\gamma^2 x/2), \quad x > 0. \tag{3.16}$$

It follows that the Lebesgue density $w(x)$ of the Lévy measure of $Z_1$ as given by (3.7) is

$$w(x) = \frac{1}{\sqrt{2\pi}} \frac{\delta}{2} \left( \frac{1}{x} + \gamma^2 \right) x^{-1/2} \, \exp(-\gamma^2 x/2), \quad x > 0. \tag{3.17}$$

The second term in the characteristic triplet vanishes. Thus, the background driving Lévy process is purely discontinuous. This follows also from the fact that all Lévy processes constructed from infinitely divisible probability distributions with support $\mathbb{R}_+$ have no Brownian part.

The BDLP $(Z_t)_{t\geq 0}$ of the IG($\delta, \gamma$) Ornstein-Uhlenbeck process $(Y_t)_{t\geq 0}$ is a sum of two independent processes $Z_t = Q_t + P_t$, where $(Q_t)_{t\geq 0}$ is an IG Lévy process with parameters $Q_1 \sim \text{IG}(\delta/2, \gamma)$ and $(P_t)_{t\geq 0}$ is of the form

$$P_t = \gamma^{-2} \sum_{i=1}^{N_t} u_i^2, \tag{3.18}$$

with Poisson process $N_t$ of rate $(\delta\gamma/2)^{-1}$ and the $u_i$ being independent standard normal and independent of the process $N_t$.

The *upper tail integral*, which is a building block for the simulation of Ornstein-Uhlenbeck type processes, is defined as

$$U(x) = \int_x^\infty w(y) dy. \tag{3.19}$$

If $Y_t \sim \text{IG}(\delta, \gamma)$ we obtain

$$U(x) = \frac{\delta}{\sqrt{2\pi}} x^{-1/2} \exp\left( -\frac{1}{2}\gamma^2 x \right). \tag{3.20}$$

This follows from the relation $U(x) = xv(x)$ of upper tail integral and Lévy measure.

### Gamma Ornstein-Uhlenbeck type Processes

From a numerical point of view the Gamma Ornstein-Uhlenbeck process is particularly interesting. It is obtained as a generalized inverse Gaussian process with parameters[3] $\lambda > 0$, $\delta = 0$ and $\gamma^2 > 0$. More exactly, $\Gamma(\lambda, \gamma^2/2) = \text{GIG}(\lambda, 0, \gamma^2)$. Here the Gamma distribution has the following density

$$\frac{(\gamma^2/2)^\lambda}{\Gamma(\lambda)} x^{\lambda-1} \exp\left( -\gamma^2 x/2 \right), \qquad x > 0. \tag{3.21}$$

---

[3]Kotz and Johnson (1983, pp. 292–298) describe a more general version of the Gamma distribution with parameters $(\vartheta, \kappa, \eta)$. The parameters applied here are related in the following way $(\vartheta, \kappa, \eta) = (2/\gamma^2, \lambda, 0)$. In the mentioned article further properties of Gamma distributions are given. Also for Gamma distributions conflicting parametrizations are in use.

Following Barndorff-Nielsen and Shephard (1999, Theorem 2.2) the Gamma distribution has the following Lévy measure

$$v(x) = \frac{\lambda}{x}\, \exp\!\big(-\gamma^2 x/2\big). \tag{3.22}$$

Consequently, the upper tail integral is given by

$$U(x) = xv(x) = \lambda \exp(-\gamma^2 x/2), \tag{3.23}$$

which we can invert analytically (in contrast to the IG case)

$$U^{-1}(x) = \max\{0;\ -2\log(x/\lambda)/\gamma^2\}. \tag{3.24}$$

The inversion of $U$ is necessary for the simulation by Rosinski expansion in Section 3.7. Since we have an analytic inversion of the upper tail integral $U$, the computation of option prices based on the simulation of integrated volatilities is 30 times faster for Gamma Ornstein-Uhlenbeck type processes than for inverse Gaussian Ornstein-Uhlenbeck type processes.

The symmetric centered Variance-Gamma distribution is constructed also as a variance-mean mixture (Madan, Carr, and Chang 1998) with respect to $(\delta, \gamma)$ and has the density

$$h(z) = \frac{\gamma^{\gamma^2+1}}{\sqrt{\lambda\pi}\,\Gamma(\gamma^2/2)2^{(\gamma^2/2)}} \left(\frac{x^2}{2\lambda}\right)^{(\gamma^2-1)/4} \mathrm{K}_{(\gamma^2-1)/2}\left(\frac{\gamma^2}{\sqrt{2\lambda}}\,x\right) \tag{3.25}$$

$$x = z - \frac{\gamma^2}{2}\ln\left(1 - \frac{2\lambda}{\gamma^4}\right), \tag{3.26}$$

where $\Gamma$ is the gamma function. This follows after some computations from Madan, Carr, and Chang (1998). Variance and kurtosis are

$$\mathrm{Var}X = \frac{2\lambda}{\gamma^2} \tag{3.27}$$

$$\mathrm{E}(X - \mathrm{E}\,X)^4 = 12\,\frac{\lambda^2}{\gamma^4}\left(\frac{2}{\gamma^2} + 1\right) \tag{3.28}$$

Of course, mean and skewness vanish in the symmetric centered case.

## Other Ornstein-Uhlenbeck type Processes

In the preceding part of this section we have introduced those Ornstein-Uhlenbeck type processes which are—from our point of view—the most interesting for financial applications. Recall, that the construction of Ornstein-Uhlenbeck type processes is possible for all self-decomposable distributions: NIG Ornstein-Uhlenbeck type processes are described in Barndorff-Nielsen (1998) and the Ornstein-Uhlenbeck type processes for other subclasses of the GIG distribution are examined in Barndorff-Nielsen and Shephard (1998). In particular, the upper tail integrals of positive hyperbolic and inverse Gamma distributions are given in Barndorff-Nielsen and Shephard (1998, Section 4.6)

## 3.3 The Model

Given a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ we consider a financial market consisting of two assets. The first one is a bank account or risk free asset $(B_t)_{0 \leq t \leq T}$ evolving according to the equation

$$
\begin{aligned}
dB_t &= rB_t dt, \quad 0 \leq t \leq T \\
B_0 &= 1.
\end{aligned}
\tag{3.29}
$$

As usual, we are not modelling the stock $(S_t)_{0 \leq t \leq T}$ but the *discounted* log-price $X_t = \log(S_t/S_0)$, which is given as a solution of the following stochastic differential equations

$$
dX_t = \sigma_t dW_t + (\mu + \beta \sigma_t^2) dt
\tag{3.30}
$$

$$
d\sigma_t^2 = -\tau \sigma_t^2 dt + dZ_{\tau t},
\tag{3.31}
$$

where $\tau > 0$ and $(Z_t)_{0 \leq t \leq T\tau}$ is the BDLP such that $\sigma^2$ is a positive, (strictly) stationary process with càdlàg path and $\sigma_t^2 \sim$ GIG for all $0 \leq t \leq T$. The processes $\sigma^2$ and $W$ are assumed to be independent and adapted to the filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$. The solution is given by

$$
X_t = \int_0^t \sigma_u dW_u + \beta \sigma_t^{2*} + \mu t, \quad 0 \leq t \leq T.
\tag{3.32}
$$

In the sequel we use the following notation for the *integrated volatility*

$$
\sigma_t^{2*} = \int_0^t \sigma_u^2 du, \quad 0 \leq t \leq T.
\tag{3.33}
$$

While $\sigma^2$ has purely discontinuous, càdlàg sample paths, the integrated volatility $\sigma^{2*}$ has continuous sample paths. By construction, the volatility process $\sigma^2$ has an autocorrelation function

$$
\mathrm{acf}(s) = \exp(-\tau s).
\tag{3.34}
$$

Note, that diffusion models often have a similar autocorrelation structure (Barndorff-Nielsen and Shephard 1998).

For $\mu = \beta = 0$ the process $X$ is a continuous local martingale and $[X]_t = \sigma^{2*}, 0 \leq t \leq T$. Barndorff-Nielsen and Shephard (1999, equation 40) give a calculus for computing all the cumulants for any weighted sum of the paths of the log-price, which offers more analytical tractability than most bivariate diffusion models posses. In particular they compute the mean and variance of $X_t$.

## 3.4 The Minimal Martingale Measure

It is well known that stochastic volatility models of the type introduced in Section 3.3 are arbitrage-free. As a consequence we get $\mathcal{M}_e \neq \emptyset$, where $\mathcal{M}_e$ denotes the set of the equivalent martingale measures, i.e. $\mathcal{M}_e = \{Q \sim P : (S_t)_{0 \leq t \leq T}$ is a Q-martingale$\}$, where $(S_t)_{0 \leq t \leq T}$ is the discounted price process given by $S_t = S_0 \exp(X_t)$, $0 \leq t \leq T$.

Following the description of the model, we define and characterize a particular martingale measure $Q$ equivalent to the real-world measure $P$. Hence, we can compute prices for a contingent claim at time $T$ given by a random variable $H \in \mathcal{L}^2(\Omega, \mathcal{F}_T, Q)$. We start with some preparatory results concerning the volatility process.

**Lemma 3.35.** *Let $(X_t)_{0 \leq t \leq T}$ be a positive, (strictly) stationary stochastic process adapted to the given filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$ and with one-dimensional marginal distribution given by*

$$X_t \sim \mathrm{GIG}(\lambda, \delta, \gamma) \quad \text{with} \quad \delta\gamma > 0. \tag{3.36}$$

*Then the following holds P-a.s.*

$$\int_0^T X_t dt < \infty \quad \text{and} \quad \int_0^T X_t^{-1} dt < \infty. \tag{3.37}$$

*Proof.* The proof follows from Liptser and Shiryayev (1977, Chapter 1); see Jørgensen (1982, p. 14) for the integrability conditions on GIG distributions. Let us denote by $Y \sim \mathrm{GIG}(\lambda, \delta, \gamma)$ the marginal distribution of the stationary process $X$. Since $\delta\gamma > 0$ we have that $\mathrm{E}[Y] < +\infty$ implying that

$$\int_0^T \mathrm{E}[X_t] dt = T \, \mathrm{E}[Y] < +\infty. \tag{3.38}$$

Applying the positivity of $X_t$ yields

$$\mathrm{E}\left[\int_0^T X_t dt\right] = \int_0^T \mathrm{E}[X_t] dt = T \, \mathrm{E}[Y], \tag{3.39}$$

and hence P-a.s.

$$\int_0^T X_t dt < \infty. \tag{3.40}$$

The same reasoning applies to the stationary stochastic process $(X_t^{-1})_{0 \leq t \leq T}$ noticing that its one-dimensional marginal distributions are given by $Y^{-1} \sim \mathrm{GIG}(-\lambda, \gamma, \delta)$ and that $\mathrm{E}[Y^{-1}] < +\infty$, being $\delta\gamma > 0$. $\qquad \square$

**Remark 3.41.** *Jørgensen (1982, p. 14) gives conditions for the finiteness of $\mathrm{E}[Y^{-1}]$ for $Y \sim \mathrm{GIG}(\lambda, \delta, \gamma)$ which are, in particular, fulfilled for Gamma distributions with $\lambda > 1$ and inverse Gamma distributions with $\lambda < 0$ (in the parametrizations as GIG distributions, see Table 1.1).*

Our project is to investigate the existence of a $Q \in \mathcal{M}_e$ which is *minimal* in some sense. In the sequel we follow the approach of Föllmer and Schweizer (1991) and Schweizer (1991) to the pricing of derivatives in incomplete markets in a slightly more general setup.

Assume that the price process $S$ is a strictly positive, continuous semi-martingale with decomposition

$$S_t = S_0 + M_t + A_t, \quad 0 \leq t \leq T,$$

where $(M_t)_{0 \leq t \leq T}$ is a local martingale and $(A_t)_{0 \leq t \leq T}$ a predictable process with bounded variation. The dynamics of $X$ are given by (3.30). Using the Itô formula we can describe the dynamics of $S$ by

$$dS_t = S_t \sigma_t dW_t + S_t \left(\mu + \left(\beta + \frac{1}{2}\right)\sigma_t^2\right) dt. \tag{3.42}$$

The integral $\left(\int_0^t S_t \sigma_t dW_t\right)_{0 \leq t \leq T}$ is a martingale since $W$ and $\sigma$ are independent and the second part in equation (3.42) is predictable and has finite variation. Hence, we have derived the Doob-Meyer decomposition of $S$.

**Definition 3.43 (Minimal Martingale Measure).** *An equivalent martingale measure $Q \in \mathcal{M}_e$ is said to be* minimal *if*

**(min1)** $Q = P$ *on $\mathcal{F}_0$;*

**(min2)** $G_t = \mathrm{E}\left[\dfrac{\mathrm{d}Q}{\mathrm{d}P}\,\Big|\,\mathcal{F}_t\right]$ *is a continuous $P$-martingale;*

**(min3)** *for any continuous $P$-local martingale $L$ which is orthogonal to $S$, in the sense that $(\langle L, S\rangle_t) \equiv 0$, the process $L$ is also a (continuous) $Q$-local martingale.*

**Proposition 3.44 (Change of Measure).** *Let $(\sigma_t)_{0 \le t \le T}$ in (3.31) be such that $P$-a.s.*

$$\int_0^T \sigma_t^{-2}\mathrm{d}t < +\infty. \tag{3.45}$$

*Then the process $(G_t)_{0 \le t \le T}$ defined by*

$$G_t = \exp\left(-\int_0^t \varphi_s\, \mathrm{d}W_s - \frac{1}{2}\int_0^t \varphi_s^2\, \mathrm{d}s\right), \tag{3.46}$$

*with*

$$\varphi_t = \frac{\mu}{\sigma_t} + \left(\beta + \frac{1}{2}\right)\sigma_t \tag{3.47}$$

*is a $P$-martingale and the probability measure $Q$ defined by*

$$\mathrm{d}Q = G_T\, \mathrm{d}P \tag{3.48}$$

*is the minimal martingale measure.*

*Proof.* See Nicolato and Prause (1999) for a detailed proof. □

## 3.5 Risk-Minimizing Hedging Strategies

To hedge against a contingent claim $H$, we use a portfolio strategy which involves the stock $S$ and the riskless bond $B = 1$ (for simplicity) and which yields the payoff $H$ at the terminal time $T$. Recall, that the asset price (only in this chapter) is discounted for simplicity.

The processes $(\xi_t)_{0 \le t \le T}$ and $(\eta_t)_{0 \le t \le T}$ denote the amounts of stock and bond, respectively, held at time $t$. We assume that $\xi$ is predictable and $\eta$ is adapted. The *value process* of the portfolio is then given by

$$V_t = \xi_t S_t + \eta_t, \quad 0 \le t \le T. \tag{3.49}$$

The *cost process* up till time $t$ is defined as

$$C_t = V_t - \int_0^t \xi_s dS_s, \quad 0 \le t \le T. \tag{3.50}$$

We only admit strategies $(\xi, \eta)$ such that the value process $(V_t)_{0 \le t \le T}$ and the cost process $(C_t)_{0 \le t \le T}$ are square-integrable, have right-continuous paths and satisfy

$$V_T = H, \quad 0 \le t \le T. \tag{3.51}$$

We also require the integrability condition

$$E\left[\int_0^T \xi_s^2 d\langle S\rangle_s + \left(\int_0^T |\xi_s|\, d|A|_s\right)^2\right] < \infty, \tag{3.52}$$

which ensures that the process of stochastic integrals in (3.50) is well defined and belongs to the space $\mathcal{S}^2$ of square integrable semimartingales. These strategies are denoted as *admissible*.

In the examined model we cannot achieve that the strategy is *self-financing*, which is defined as

$$C_t = C_T = H, \quad 0 \le t \le T. \tag{3.53}$$

Therefore, we look at strategies based on *risk-minimization* approaches. Föllmer and Sondermann (1986) proposed latter approach to find an equivalent martingale measure in an incomplete market. We are going to single out that admissible strategy which minimizes the remaining risk

$$E\left[(C_T - C_t)^2 | \mathcal{F}_t\right] \tag{3.54}$$

for each time point $t < T$. Risk-minimizing strategies are typically no longer self-financing, but they are *mean-self-financing*

$$E[C_T - C_t | \mathcal{F}_t] = 0, \quad 0 \le t \le T, \tag{3.55}$$

i.e. the cost process $C$ is a martingale.

First, let us look at the case where $S$ is already a martingale. Then the Kunita-Watanabe decompositon holds

$$H = H_0 + \int_0^T \xi_s^H dS_s + L_T^H, \tag{3.56}$$

with $H_0 \in \mathcal{L}^2(\Omega, \mathcal{F}_0, P)$ and $L_H$ a square-integrable martingale orthogonal to $S$. The risk minimizing strategy is in this case given by

$$\xi := \xi^H, \quad \eta := V - \xi \cdot S, \tag{3.57}$$

with value process

$$V_t := H_0 + \int_0^t \xi_s^H dS_s + L_t^H, \quad 0 \le t \le T. \tag{3.58}$$

However, in general the price process is not a martingale.

To single out a particular martingale measure in a market described by semimartingales, Schweizer (1988) introduced the concept of *local risk-minimization*. A strategy is called locally risk-minimizing if the remaining risk $E[(C_T - C_t)^2 | \mathcal{F}_t]$ for any $t < T$ is minimal under all infinitesimal perturbations of the strategy at time $t$ (see Schweizer (1991, 1992) for exact definitions and further results). This property is essentially equivalent to the following criterium which we will apply in the sequel.

**Definition 3.59.** *An admissible strategy $(\xi, \eta)$ is called* optimal *if the associated cost process $C$ is a square-integrable martingale which is orthogonal to $M$ under $P$.*

**Proposition 3.60.** *The existence of an optimal strategy is equivalent to a decomposition*

$$H = H_0 + \int_0^T \xi_s^H dS_s + L_T^H, \tag{3.61}$$

*with $H_0 \in \mathcal{L}^2(\Omega, \mathcal{F}_0, P)$, where $\xi^H$ satisfies the integrability condition in Föllmer and Schweizer (1991, equation (2.8)) and $L^H$ is a square-integrable martingale orthogonal to $M$. For such a decomposition the associated optimal strategy $(\xi, \eta)$ is given by (3.57) and (3.58).*

The existence of the optimal hedging strategy follows with the next theorem from the existence of the minimal martingale measure in the sense of Föllmer and Schweizer (1991). Unfortunately, we have not shown that the density process in the Girsanov transformation and the price process are square integrable in the IGOU model. Therefore, the following theorem and the explicit computation of the hedging strategy are not true in general.

Nevertheless, we state the results since we think that it should be possible to show that both processes are sqare integrable, possibly only for a subset of parameters.[4]

**Theorem 3.62 (Optimal Strategy and Minimal Martingale Measure).** *The optimal strategy, hence also the corresponding decomposition* (3.61) *is uniquely determined. It can be computed in terms of the minimal martingale measure $Q$ given in the sense of Föllmer and Schweizer (1991): If $(V_t)_{0 \le t \le T}$ denotes a right-continuous version of the martingale*

$$V_t := \mathrm{E}^Q[H | \mathcal{F}_t], \quad 0 \le t \le T, \tag{3.63}$$

*then the optimal strategy $(\xi, \eta)$ is given by* (3.57)*, where*

$$\xi^H = \frac{d\langle V, S \rangle}{d\langle S \rangle} \tag{3.64}$$

*is obtained by projecting the $Q$-martingale $V$ on the $Q$-martingale $S$.*

Föllmer and Schweizer (1991, Theorem 3.14) proved this theorem.

## Explicit Computation of the Hedging Strategy

For the explicit computation of the locally risk-minimizing hedging strategy we follow Musiela and Rutkowski (1997, Section 5.1.3). Assume, that the option price $V$ follows

$$V_t = v(S_t, t) \tag{3.65}$$

for some function $v : \mathrm{I\!R} \times [0, T] \to \mathrm{I\!R}$. The optimal replicating strategy is then given by the following functions

$$(\xi, \eta) = \big(g(S_t, t), h(S_t)\big), \tag{3.66}$$

where $g : \mathrm{I\!R} \times [0, T] \to \mathrm{I\!R}$ and $h : \mathrm{I\!R} \times [0, T] \to \mathrm{I\!R}$ are functions unknown at the moment. The wealth process equals then

$$V_t(\xi, \eta) = g(S_t, t)S_t + h(S_t, t)B_t = v(S_t, t). \tag{3.67}$$

---

[4]Results in this direction will be included in a final version of Nicolato and Prause (1999).

With property (3.61) we obtain the following dynamics of the process $(V_t(\xi, \eta))_{0 \leq t \leq T}$

$$dV_t(\xi, \eta) = g(S_t, t)dS_t + h(S_t, t)dB_t = \int \xi^H \, dS_t + \underbrace{H_0 + L_t^H}_{=: \, h(S_t, t)dB_t}. \tag{3.68}$$

Now we plug in the dynamics of $S$ given by (3.42)

$$dV_t(\xi, \eta) = g(S_t, t)S_t\sigma_t dW_t + g(S_t, t)S_t\Big(\mu + \Big(\beta + \frac{1}{2}\Big)\sigma_t^2\Big)dt + h(S_t, t)r \, dt. \tag{3.69}$$

From (3.67) we obtain

$$\eta = h(S_t, t) = B_t^{-1}\big(v(S_t, t) - g(S_t, t)\big). \tag{3.70}$$

We shall search for the wealth function $v$ in the class of smooth functions on the open domain $\mathcal{D} = (0, \infty) \times (0, T)$, more exactly we assume that $v \in C^{2,1}(\mathcal{D})$. Now we apply the Itô formula (in the formulation of Musiela and Rutkowski (1997, p. 463)) conditionally on $\sigma$. This is possible since the driving Brownian motion $W$ and the volatility process $\sigma^2$ are mutually independent. See Section 3.6 for another application of this argument. As usual, the subscripts to $v$ denote the corresponding partial derivatives.

$$\begin{aligned} dv(S_t, t) &= v_t(S_t, t)dt + S_t\Big(\mu + \Big(\beta + \frac{1}{2}\Big)\sigma_t^2\Big)v_s(S_t, t)dt \\ &\quad + S_t\sigma_t v_s(S_t, t)dW_t + \frac{1}{2}S_t^2\sigma_t^2 v_{ss}(S_t, t)dt \end{aligned} \tag{3.71}$$

Now we consider the process $Y_t = v(S_t, t) - V_t(\xi, \eta)$, $0 \leq t \leq T$. Then we obtain for the dynamics of $(Y_t)_{0 \leq t \leq T}$

$$\begin{aligned} dY_t &= dv(S_t, t) - dV_t(\xi, \eta) \\ &= v_t(S_t, t)dt + S_t\Big(\mu + \Big(\beta + \frac{1}{2}\Big)\sigma_t^2\Big)v_s(S_t, t)dt \\ &\quad + S_t\sigma_t v_s(S_t, t)dW_t + \frac{1}{2}S_t^2\sigma_t^2 v_{ss}(S_t, t) \\ &\quad - g(S_t, t)S_t, \sigma_t dW_t - g(S_t, t)S_t\Big(\mu + \Big(\beta + \frac{1}{2}\Big)\sigma_t^2\Big)dt \\ &\quad - h(S_t, t)r \, dt. \end{aligned} \tag{3.72}$$

Since $V_t(\xi, \eta) = v(S_t, t)$ holds, $Y$ vanishes identically. Thus, the diffusion term in (3.72) has to be zero. We obtain

$$\int S_u\sigma_u\big(v_s(S_t, t) - g(S_t, t)\big)dW_t = 0 \tag{3.73}$$

and hence

$$v_s(S_t, t) = g(S_t, t) \quad \forall(s, t) \in \mathbb{R}_+ \times [0, T]. \tag{3.74}$$

Consequently, the hedging strategy, conditional on $\sigma$, is given by the derivative $v_s$ of the option price w.r.t. the stock price. Recall, that this is a locally risk-minimizing heging strategy only under the condition that the density process and the price process are square integrable.

## 3.6 Conditional Log-Normality

The calculation of IGOU prices can be split in two parts. We use th indepencende of $\sigma^2$ and $W$ to prove that $X_T$ has a log-normal distributon conditionally on the integrated volatility $\sigma_T^{2*}$ at maturity.

**Theorem 3.75.** *In the stochastic volatility model introduced in Section 3.3 and the measure transformation from Theorem 3.44 we can compute prices for a sufficiently regular terminal payoff function h by*

$$\mathrm{E}^Q[h(S_T)] = \int_{\mathbb{R}_+} \mathrm{BS}(h, S_0, v, T) \lambda^Q(dv). \tag{3.76}$$

*Here $\lambda^Q$ denotes the distribution of the average volatility $v = \sqrt{\sigma_T^{2*}/T}$ at time $T$ under $Q$ and $\mathrm{BS}(h, S_0, v, T)$ the Black-Scholes price of an option with terminal payoff function h.*

Compare latter theorem with similar results in Frey (1996, Chapter 6.2.1), Hull and White (1987).

*Proof.* Under $P$ we have that $\sigma$ and $W$ are independent. Therefore, we could write the original probability space as a product space

$$(\Omega, \mathcal{F}, P) = \left((\Omega, \mathcal{F}^\sigma) \times (\Omega, \mathcal{F}^W), P^\sigma \otimes P^W\right) \tag{3.77}$$

where $\mathcal{F}^\sigma, P^\sigma$ and $\mathcal{F}^W, P^W$ denote the filtrations and the distributions associated with the processes $\sigma$ and $W$ respectively. Consequently, we can condition on the filtration $\mathcal{F}^\sigma$ and the process $W$ remains conditionally on $\mathcal{F}^\sigma$ a Brownian motion. For a fixed $v \in \Omega^\sigma$ the following holds

$$\mathrm{E}^Q[h(S_T)|\sigma = v] = \mathrm{BS}(h, S_0, v, T). \tag{3.78}$$

The measure transformation $P \to Q$ in Theorem 3.44 for a given $v$ is equal to the measure transformation in the Black-Scholes model. Consequently,

$$\begin{aligned}
\mathrm{E}^Q[h(S_T)] &= \mathrm{E}^Q\left[\mathrm{E}[h(S_T) \,|\, \sigma = v]\right] \\
&= \mathrm{E}^Q[\mathrm{BS}(h, S_0, v, T)] = \int_0^\infty \mathrm{BS}(h, S_0, v, T) \lambda^Q(dv),
\end{aligned}$$

where $\lambda^Q$ denotes the distribution of the average volatility. $\qquad \square$

**Remark 3.79.** *For the actual computation we can use that $\lambda^Q$ is equal to the distribution of the integrated volatility $\lambda^P$ under $P$. This follows from the fact, that we have not changed the law of the volatility process $\sigma$ in Girsanov transformation.*

**Corollary 3.80 (European Call Option).** *A European call option, i.e. a contingent claim with terminal payoff function $(S_T - K)^+$, has the IGOU price*

$$C_{\mathrm{IGOU}} = \int_0^\infty C_{\mathrm{BS}}(S_0, K, v, T) \lambda^P(v),$$

81

where $C_{\mathrm{BS}}$ denotes the Black-Scholes formula for European call options

$$C_{\mathrm{BS}}(S_0, K, v, T) = S\,\Phi(x) - e^{-rT} K\,\Phi\big(x - \sigma\sqrt{T}\,\big)$$
$$x = \frac{\ln(S\,e^{rT}/K)}{\sigma\sqrt{T}} + \frac{\sigma\,\sqrt{T}}{2}$$

and $\Phi$ is the cumulative density function of a standard normal distribution.

An extension of this result to Ornstein-Uhlenbeck type models incorporating leverage effects (Barndorff-Nielsen and Shephard 1999) is not possible with the same argument, since the Brownian motion $W$ and the instantaneous volatility $\sigma$ are no longer independent.

## 3.7 Computation of Prices by Rosinski Expansion

In the previous section we have shown that *in principle* it is possible to perform arbitrage-pricing whenever the one-dimensional marginal distibution of the stationary process $\sigma^2$ is chosen within the class of the $\mathrm{GIG}(\lambda, \delta, \gamma)$ laws with $\delta\gamma > 0$ or fulfills the conditions given in Remark 3.41. In order to show that the computation of formula (3.76) is feasible *in practice* we restrict ourselves to the case of $\sigma_t^2 \sim \mathrm{IG}(\delta, \gamma)$ with $\delta\gamma > 0$. The crucial point in the computation of prices by formula (3.76) is to determine the distribution $\lambda^P$ of the integrated volatility $\sigma_T^{*2}$. In the sequel we follow Barndorff-Nielsen and Shephard (1998, Section 4.5).

Using the property of the integrated volatility that

$$\sigma_t^{2*} = \frac{1}{\tau}\big(Z_{\tau t} - \sigma_t^2 + \sigma_0^2\big), \tag{3.81}$$

the distribution $\lambda^P$ of the integrated volatility $\sigma^{2*}$ may be simulated. A direct computation of $\sigma_T^{2*}$ using (3.81) is not possible, since $(Z_{\tau t})_{0 \leq \tau t \leq T}$ and $(\sigma_t^2)_{0 \leq t \leq T}$ are not independent.

The equation above implies for simulated increments of the integrated volatility $\bar{\sigma}_{\Delta(n+1)}^2 := \sigma_{\Delta(n+1)}^{2*} - \sigma_{\Delta n}^{2*}$ the following identity

$$\bar{\sigma}_{n+1}^2 = \frac{1}{\tau}\left(Z_{\tau\Delta(n+1)} - Z_{\tau\Delta n} + \sigma_{\Delta n}^2 - \sigma_{\Delta(n+1)}^2\right), \tag{3.82}$$

with sampling interval $\Delta > 0$. We can simulate sequences of integrated volatilities by simulating the bivariate process of the BDLP $Z$ and the instantaneous volatility $\sigma^2$ in the following way

$$\sigma_{\Delta(n+1)}^2 = e^{-\tau\Delta}\sigma_{\Delta n}^2 + \underbrace{e^{-\tau\Delta}\int_{\tau\Delta n}^{\tau\Delta(n+1)} e^{-(\tau\Delta(n+1)-s)}dZ_s}_{=:\,w_{1,n+1}}, \tag{3.83}$$

$$Z_{\tau\Delta(n+1)} = Z_{\tau\Delta n} + \underbrace{\int_{\tau\Delta n}^{\tau\Delta(n+1)} dZ_s}_{=:\,w_{2,n+1}}.$$

82

Equation (3.83) is obtained by the application of (3.13). Using the fact that $Z$ is a Lévy process we can express $w_n$ by

$$w_n \stackrel{\mathcal{L}}{=} \begin{pmatrix} \exp(-\tau\Delta) \int_0^{\tau\Delta} e^s \, dZ_s \\ \int_0^{\tau\Delta} dZ_s \end{pmatrix}. \tag{3.84}$$

The following expansion allows us to simulate the innovations in a comfortable way. For an integrable function $f : [0, x] \to \mathbb{R}_+$ holds (Rosinski 1991)

$$\int_0^x f(s) dZ_s \stackrel{\mathcal{L}}{=} \sum_{i=1}^{\infty} U^{-1}(a_i^*/x) f(x r_i), \tag{3.85}$$

where $\{a_i^*\}$ and $\{r_i\}$ are two independent sequences of random variates with the $r_i$ independent uniform random variables on $[0, 1]$ and $a_1^* < \cdots < a_i^* < \ldots$ arrival times of a Poisson process with intensity 1. Here $U$ is the upper tail integral with the inverse

$$U^{-1}(x) = \inf\{y > 0 : U(y) \le x\}. \tag{3.86}$$

The values for $U$ in the case of IG und Gamma Ornstein-Uhlenbeck type processes are given in Section 3.2. Note, that $a_i^* - a_{i-1}^* \stackrel{\mathcal{L}}{=} -\log(u_i)$, where $\{u_i\}$ are uniformly distributed on $[0, 1]$. Furthermore, $U^{-1}$ is inverted in (3.85) for increasing $a_i^*/\tau$. Therefore, $a_i^*/\tau$ may be used as an initial upper bracket in the numerical inversion algorithm. We have used the Illinois algorithm leading to a fast convergence.

Taken together, we can simulate the increments of the integrated volatility in the following way

$$\bar{\sigma}_{n+1}^2 = \tau^{-1} \left( w_{2,n+1} + \sigma_{\Delta n}^2 - \exp(-\tau\Delta)\sigma_{\Delta n}^2 - w_{1,n+1} \right) \tag{3.87}$$

$$w_{1,n+1} = \exp(-\tau\Delta) \sum_{i=1}^{\infty} U^{-1}(a_i^*/(\tau\Delta)) \exp(\tau\Delta r_i) \tag{3.88}$$

$$w_{2,n+1} = \sum_{i=1}^{\infty} U^{-1}(a_i^*/(\tau\Delta)). \tag{3.89}$$

The simulation of the integrated volatility needs the major part of the computation time of Ornstein-Uhlenbeck type option prices. Therefore, various approaches to increase the speed of computations have been proposed, e.g. to memorize the values of the inverted upper tail integral $U^{-1}$ in a matrix. We have not applied these approaches, since in practical applications the integrated volatility of the IGOU process has to be rescaled by changing up-to-date volatility estimates. This would lead to an unpleasant data handling of the various matrices.

## 3.8 Saddle-Point Approximation to the Option Price

To apply the saddle-point approximation already proposed in Section 1.14 we have to know some results concerning the cumulant generating function of the (discounted) log-price $X_T$ at maturity in the IGOU model. The cumulant generating function, defined as $K\{z \ddagger X_T\} := \log \mathrm{E}[\exp(z X_T)]$, is quite complicated, but explicit.

**Proposition 3.90 (IGOU Cumulant Generating Function).** *The cumulant generating function* $K\{z \ddagger X_T\}$ *is finite in the interval* $(-a_-, a_+)$ *where*

$$a_- = -\frac{1 - [1 + 4\gamma^2\tau(1 - e^{-\tau T})^{-1}]^{1/2}}{2},$$

$$a_+ = \frac{1 + [1 + 4\gamma^2\tau(1 - e^{-\tau T})^{-1}]^{1/2}}{2}.$$

*It is given by* $K\{z \ddagger X_T\} = C_1(z) + C_2(z)$, *where*

$$C_1(z) = \delta\gamma\left[1 - \left(1 - \frac{(1 - e^{-\tau T})}{\gamma^2\tau}(z^2 - z)\right)^{1/2}\right],$$

$$C_2(z) = F_1(z) + F_2(z)\,\mathbf{1}_{[-b_-,b_+]}(z) + F_3(z)\,\mathbf{1}_{[-a_-,a_+]\setminus[-b_-,b_+]}(z)$$

*and the real numbers* $b_-, b_+$ *are given by*

$$b_- = -\frac{1 - [1 + 4\gamma^2\tau]^{1/2}}{2}, \quad and \quad b_+ = \frac{1 + [1 + 4\gamma^2\tau]^{1/2}}{2}.$$

*We use the following abbreviations*

$$g_1(z) = \sqrt{\gamma^2 - \tau^{-1}(z^2 - z)(1 - e^{-\tau T})},$$

$$g_2(z) = \sqrt{\gamma^2 - \tau^{-1}(z^2 - z)},$$

$$g_3(z) = \sqrt{-(\gamma^2 - \tau^{-1}(z^2 - z))},$$

*then the expressions for* $F_1(z)$, $F_2(z)$ *and* $F_3(z)$ *are given by*

$$F_1(z) = \delta(g_1(z) - \gamma),$$

$$F_2(z) = \frac{\delta(z^2 - z)}{2\tau g_2(z)}\log\left(\frac{(\gamma - g_2(z))(g_1(z) + g_2(z))}{(\gamma + g_2(z))(g_1(z) - g_2(z))}\right),$$

$$F_3(z) = \frac{\delta(z^2 - z)}{2\tau g_3(z)}\left[\arctan\left(\frac{\gamma}{g_3(z)}\right) - \arctan\left(\frac{g_1(z)}{g_3(z)}\right)\right].$$

*Moreover, the Laplace transform of* $X_T$ *is integrable and the exponential family (in the sense of Jensen 1995, p. 6) is steep.*

*Proof.* See Nicolato and Prause (1999) for the details of the proof. $\square$

Note, that the condition $a_+ > 1$ is satisfied for any values of the parameters $\tau$ and $\gamma$, enabling us to use equation (1.66) to express the price of European *put* options

$$\text{Price} = Ke^{-rT}Q\left[X_T < k - rT\right] - S_0 e^{\kappa(1)}P_1\left[X_T < k - rT\right], \tag{3.91}$$

where $\kappa(z) = K\{z \ddagger X_T\}$ and $k = \log(K/S_0)$. To apply the saddle-point approximation we have to compute the first and second derivative of the cumulant generating functions of $\kappa(z) = K\{z \ddagger X_T\}$. We look at the representation given in Propositon 3.90 and use the abbreviation

$$A := \frac{1 - e^{-\tau T}}{\gamma^2\tau}.$$

Then we get for the derivatives of $C_1$

$$C_1'(z) = \delta\gamma\frac{A}{2}\frac{2z-1}{\sqrt{1-A(z^2-z)}},$$

$$C_1''(z) = \frac{\delta\gamma A}{\sqrt{1-A(z^2-z)}} + \frac{\delta\gamma A^2(2z-1)^2}{4(1-A(z^2-z))^{3/2}}.$$

We write $B = 1 - \exp(-\tau T)$ for short and obtain

$$F_1'(z) = -\frac{\delta B}{2\tau}\frac{2z-1}{g_1(z)},$$

$$F_1''(z) = -\frac{\delta B}{\tau g_1(z)} - \frac{\delta B^2(2z-1)^2}{4\tau^2(g_1(z))^3}.$$

In the actual implementation we differentiate the functions $F_2$ and $F_3$ numerically and for numerical reasons we compute $g_1 - g_2$ as

$$g_1 - g_2 = \frac{(z^2-z)e^{-\tau T}}{\tau(g_1+g_2)}. \tag{3.92}$$

For typical estimates of IG distributions and values of $k - rT$ the CDF does exist. Moreover, $b_-$ and $b_+$ are close to $a_-$ and $a_+$ respectively. Therefore, in most cases only $F_3$ needs to be calculated.

With Proposition 3.8 all necessary conditions are fulfilled to apply Theorem 1.69 (Lugannani and Rice 1980) to approximate IGOU option prices. In Section 3.11 we will compare the numerical results of the saddle-point approximation and of the pricing by the simulation of the integrated volatility.

Note, that a better approximation may be achieved using the higher order terms given in the original paper of Lugannani and Rice (1980). In general, simulations studies indicate that the use of the first order term leads to good results (Daniels 1987).

## 3.9  Calibration of the IGOU Model

The estimation of parameters of the Ornstein-Uhlenbeck type process proceeds in two steps: First, we compute under the assumption of independence the likelihood estimate of the marginal NIG distribution using the algorithm described in Section 1.2. Then we can derive from (1.6) the mixing $IG(\delta,\gamma)$ distribution which will be the stationary distribution of the Ornstein-Uhlenbeck type volatility process. In particular, $\delta$ is a common parameter of GH and GIG distributions whereas $\gamma = \sqrt{\alpha^2 - \beta^2}$. Secondly, the autocorrelation parameter $\tau$ is estimated by regression

$$\sum_{t=0}^{T}\left|\widehat{\mathrm{acf}}(t) - \exp(-\tau t)\right|^h, \tag{3.93}$$

where $h = 1, 2$ and $\widehat{\mathrm{acf}}$ denotes the empirical ACF of the absolute returns. As an alternative the regression may be based on the cumulative ACF.

One could apply likelihood estimation not only for the stationary distribution, but also for the autocorrelation described by $\tau > 0$. Barndorff-Nielsen (1998, Section 3.2) proposed to estimate the parameters in a simulation-based likelihood approach; see also Barndorff-Nielsen and Shephard (1999, Section 5) for other estimation methods.
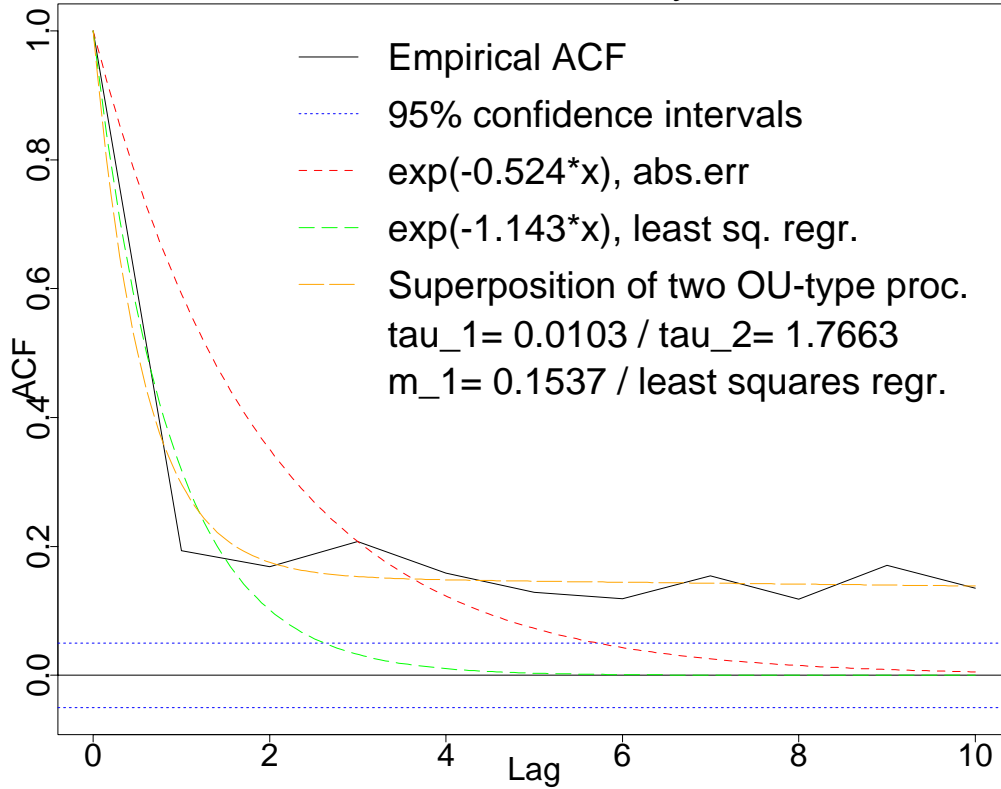
Figure 3.3: *Empirical autocorrelation function of the absolute returns and fitted autocorrelation functions of Ornstein-Uhlenbeck type and a superposition of two Ornstein-Uhlenbeck type processes.*

## 3.10 Superposition of IGOU Processes

A further refinement of the model described above is achieved by replacing the simple IG Ornstein-Uhlenbeck type process by a sum of IG processes (denoted as supIGOU; see Barndorff-Nielsen (1998, Chapter 6) and Barndorff-Nielsen (1999) for other results on superpositions). We use the fact that IG distributions are closed under convolution

$$\text{IG}\Big( \sum_{i=1}^{n} \delta_i, \gamma \Big) = \underset{i=1}{\overset{n}{\ast}} \ \text{IG}(\delta_i, \gamma), \tag{3.94}$$

where $\gamma > 0$ and $\delta_i > 0$ for $i = 1, \ldots, n$. Assume, $Y^i$, $i = 1, \ldots, n$ are independent and stationary IG Ornstein-Uhlenbeck type processes with parameters $(\delta_i, \gamma, \tau_i)$. Then $Y = \sum_{i=1}^{n} Y^i$ is a stochastic process with $\text{IG}(\sum_{i=1}^{n} \delta_i, \gamma)$ one-dimensional stationary distribution and autocorrelation function

$$\text{acf}(s) = \sum_{i=1}^{n} \exp(-s\tau_i) \, \delta_i/\delta, \tag{3.95}$$

where $\delta = \sum_{i=1}^{n} \delta_i$.

## Estimation

With the same two-step approach as described in Section 3.9 it is also possible to estimate parameters in the case of a superposition of $n$ IG Ornstein-Uhlenbeck type processes. Again, from the NIG distribution we get the mixing $IG(\delta, \gamma)$ distribution and we will use this distribution as the IG law of the superposition.

The next step is the estimation of the autocorrelation and weight parameters $(\lambda_1, \ldots, \lambda_n, m_1, \ldots, m_n)$. To obtain stable numerical results one could minimize

$$\int_0^\infty \left( \widehat{\mathrm{acf}}(t) - \sum_{j=1}^n m_j \exp(-\tau_j t) \right)^2 dt \approx \sum_{i=0}^{[T/\Delta]} \left( \widehat{\mathrm{acf}}(i\Delta) - \sum_{j=1}^n m_j \exp(-\tau_j i\Delta) \right)^2,$$

with $\sum_{j=1}^n m_j = 1$. We choose $\Delta = 1/2$ and interpolate the empirical autocorrelation function $\widehat{\mathrm{acf}}(t)$ linearly between the integer values. The parameters of the $n$ IGOU processes are now given as $(\delta m_j, \gamma, \tau_j)$, $j = 1, \ldots, n$. Hence, the superposition has the $IG(\delta, \gamma)$ distribution obtained from the mixture representation.

We restrict ourselves to a superposition of two IG Ornstein-Uhlenbeck type processes. The estimation procedure described above yields the estimates $(\delta, \gamma)$ for the IG distribution and the parameters $(\tau_1, \tau_2, m_1, m_2)$ with $m_1 + m_2 = 1$. Now we can construct two independent IGOU processes with parameters $(\delta m_i, \gamma, \tau_i)$, $i = 1, 2$. Their superposition has an $IG(\delta, \gamma)$ marginal density and the desired autocorrelation structure

$$\mathrm{acf}(s) = m_1 \exp(-\tau_1 s) + (1 - m_1) \exp(-\tau_2 s). \tag{3.96}$$

In Figure 3.3 we present the estimation results for the autocorrelation functions of the IGOU and supIGOU processes for Bayer stock returns. Obviously, simple IG Ornstein-Uhlenbeck type processes have not enough flexibility to fit empirical autocorrelation functions especially for increasing time lags. However, a superposition of only two IG Ornstein-Uhlenbeck type processes allows to model the dependence structure and the long-range dependence precisely. Estimated supIGOU processes usually consist of two distinct components: One IGOU process with substantial long-range dependence but a lower weight (e.g. in the case of Bayer stocks we have $\tau_1 = 0.0103$ and $m_1 = 0.1537$) and a second IGOU process with a much faster decreasing autocorrelation function.

To simulate the distribution of the integrated volatility $\lambda^P$ of the superposition we have to simulate the integrated volatilities $\lambda_j^P, j = 1, 2$ of both independent Ornstein-Uhlenbeck type processes and obtain

$$\lambda^P = \lambda_1^P * \lambda_2^P. \tag{3.97}$$

To get a single simulation result of (3.97) we simulate independently paths of both IGOU processes and sum the integrated volatilities with weights $m_i$. The repetition of this procedure approximates the distribution of the integrated volatility $\lambda^P$ of the superposition.

A statistical motivation of this two-step approach, which was already used in Section 3.9, is given by Jiang and Pedersen (1998). They examine the parameter estimation for a superposition of two IG Ornstein-Uhlenbeck type processes as an unobserved volatility process and prove a CLT for marginal NIG distributions in the case of an increasing observation period $T_n$ and decreasing sampling interval $\Delta_n$, i.e. for $\Delta_n \to 0$ and $T_n \to \infty$ such that $T_n \Delta_n \to 0$ for $n \to \infty$. Finally, they propose to fit the autocorrelation and weight parameters $m_1, \tau_1, \tau_2$ by regression.

87

## 3.11   Numerical Results

In this section the computational methods proposed in the Sections 3.7–3.8 are carried out for the Bayer stock estimates. The mixing IG distribution is given (in the different parametrizations of the Chapters 1 and 3) by

$$\chi = 0.0001068 \qquad \psi = 6653.032,$$
$$\delta = 0.01034 \qquad \gamma = 81.57.$$

### Simulation of Integrated Volatilities

The first step in the computation of IGOU prices is the simulaton of $\lambda^P$ as described in Section 3.7. As starting value $\sigma_0^2$ we use the mean of the IG distribution. Other starting values are possible, e.g. the current volatility estimate or a random variate sampled out of the stationary distribution. However, the latter approach does not change the IGOU option prices substantially. We realized the algorithm with $\Delta = 1$ (day) and simulated 1000 or more paths of the instantaneous volatility for each price computation. A large number of simulations is necessary to obtain stable results for IGOU prices. Therefore, the computation of a single price is rather slow.

In contrast to saddle-point approximation of option prices, the simulation of $\sigma_T^{2*}$ yields also the paths of instantaneous volatilities. Figure 3.4 shows the path of an IGOU process and the marginal distribution of the path obtained by the simulation, which is close to the stationary IG distribution of the process. The paths show *volatility clustering*, i.e. the instantaneous volatility jumps upwards and the reversion effect described by $\tau$ forces $\sigma^2$ back to smaller values. The two plots below show the volatility measured in annualized standard deviations. Clearly, the values of $\sigma^2$ belong to the usually observed range of (annualized) volatilities.

Figure 3.5 shows the paths of a superposition and of the component with long-range dependence, both measured in annualized standard deviations. We have simulated the IGOU process with $\tau = 0.01$ over a longer time period. Obviously, this component may explain significant jumps in the volatility and a slow calming down in the stock market afterwards.

### Stability of IGOU Price Simulations

How stable is the calculation of the IGOU prices, i.e. how many paths of the integrated volatility do we have to simulate? To examine this question, we have simulated the integrated volatility for various sample sizes. The simulated density of $\lambda^Q$, which is plotted in Figure 3.6 becomes sufficiently stable for at least 1000 simulations. Considering only the densities and log-densities in Figure 3.6, one would expect stable results for about 1000 simulations. However, looking at the mean values of the integrated volatilites and the prices computed for increasing sample sizes plotted in Figure 3.7, shows that the simulation error is not a quantité negligable, but it is small enough to distinguish IGOU prices from Black-Scholes prices. This is also shown in Figure 3.7.

Unfortunately, the large simulation error prevents the computation of implicit volatilities in this model. A possible by-pass would be to apply for each computation of IGOU price in the inversion algorithm the same pseudo-random-numbers. Hence, one would enforce artificially the stability of the simulation, but this would not remove the simulation error itself.
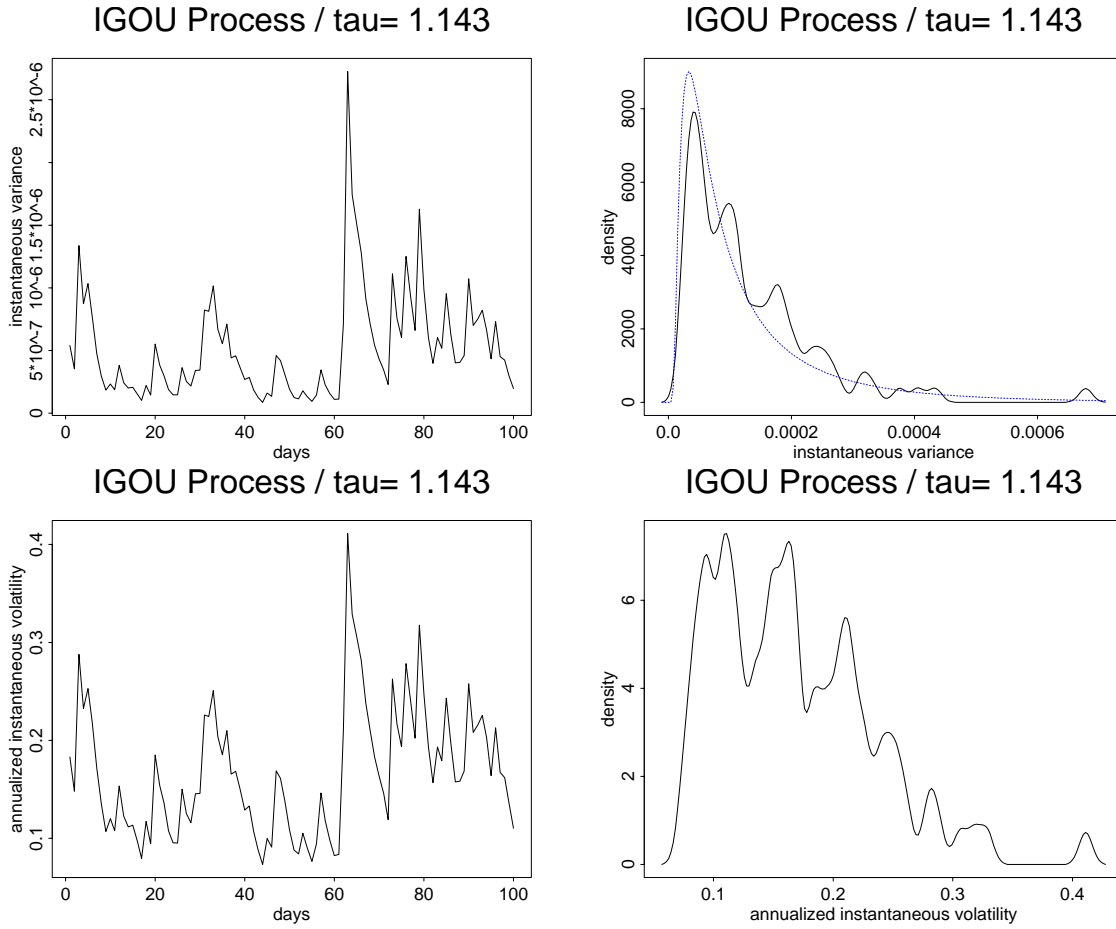
Figure 3.4: *Simulation of instantaneous volatility, 100 trading days, Bayer parameter, $\tau = 1.143$. The both graphs below (in this figure) contain the same path given as annualized volatilities.*
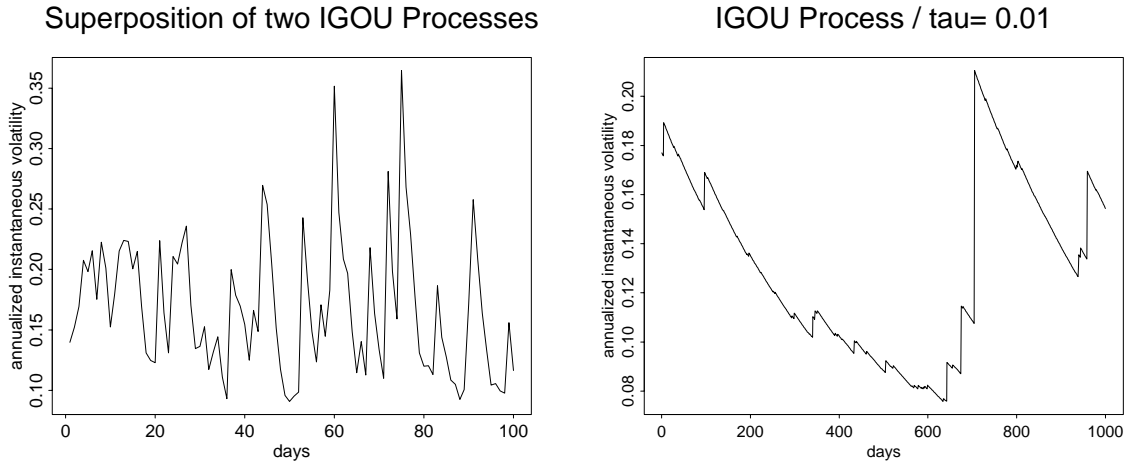


Figure 3.5: *Simulation of the (annualized) instantaneous volatility of a supIGOU process and an IGOU process with long-range dependence, i.e. small $\tau$ similar to $\tau_1$ in the supIGOU process. Left: Superposition of two IG Ornstein-Uhlenbeck type processes with $\tau_1 = 0.0103$, $\tau_2 = 1.7663$ and $m_1 = 0.1537$. Right: IG Ornstein-Uhlenbeck type process with $\tau = 0.01$. The parameters are estimates for Bayer stocks 1988–94.*
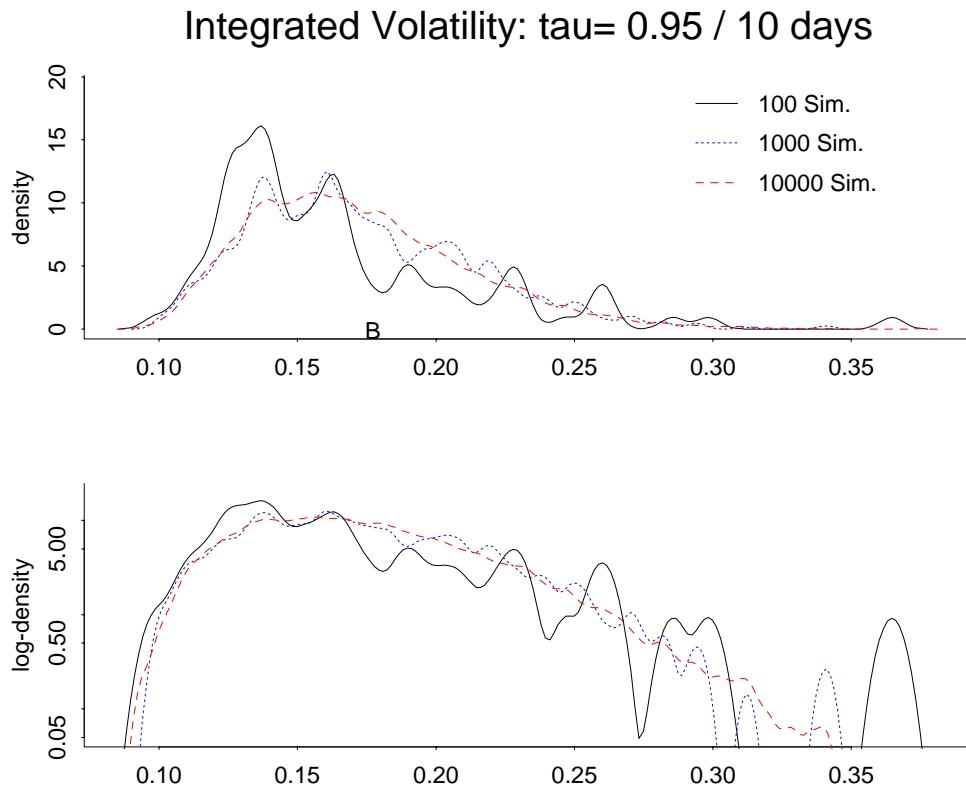
89

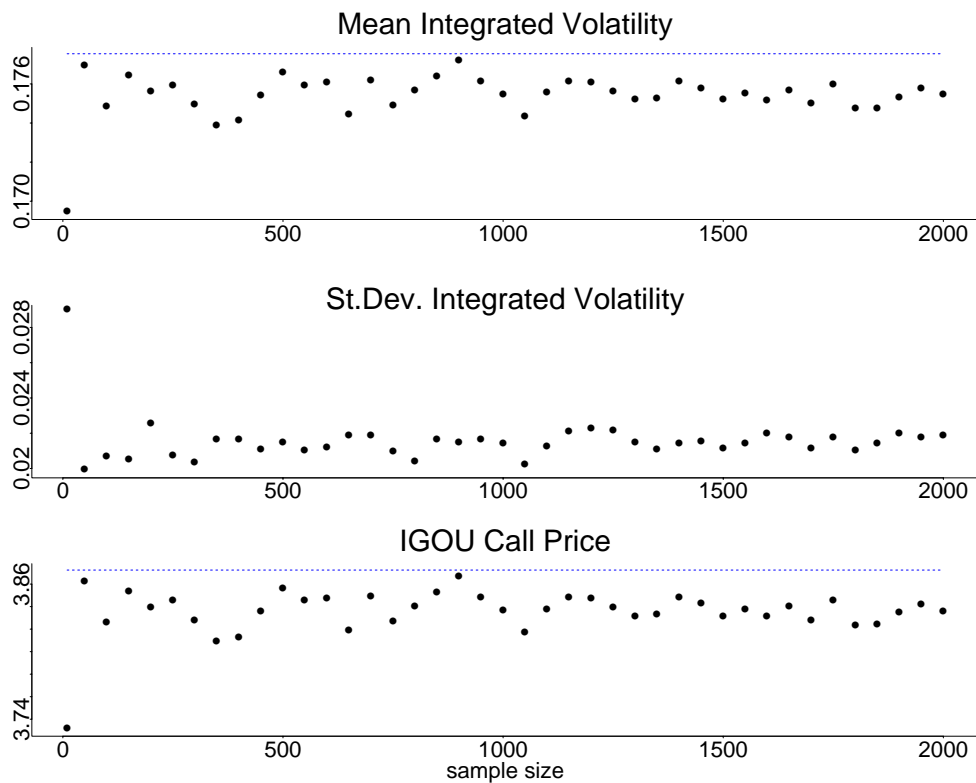Figure 3.6: *Simulation of the integrated volatility.*



Figure 3.7: *Simulation of the integrated volatility with increasing sample sizes ($\tau = 0.95$ and 50 days to expiration).*

Although it is not really satisfying, we compute prices based on 1000 simulated paths. This choice is motivated by the desire to compute the prices in a reasonable time. Finally, saddlepoint approximations will turn out to be a reasonable, less time-consuming alternative.

## Comparison of IGOU, supIGOU and Black-Scholes Prices

The first step in the evaluation of a new pricing model for derivatives—before looking at real option data—is usually the simulation of prices. Since the Black-Scholes model is the industrial standard in Finance, we could compare IGOU and Black-Scholes prices instead of plotting the IGOU prices themselves.

The IGOU prices are based on the simulation of the integrated volatility. Note, that we have plotted in Figures 3.8 and 3.9 the differences to the Black-Scholes model—which are small in comparison to the prices. As a consequence of the simulation error, the graphs are rather rough. We vary the autocorrelation parameter $\tau$ and the kurtosis of the NIG distribution to explain the effect of dependence structure and heavy tails, respectively.

The first observation is that we observe again the W-shape in the price differences to Black-Scholes prices which was already observed by Eberlein and Keller (1995) in the hyperbolic model. Obviously, prices of derivatives in the IGOU model also have the potential to explain the smile effect. We have not calculated implied volatilies for the IGOU model, but we propose an "implied" approach in Section 3.13.

In contrast to the prices in the GH Lévy model we observe an increasing difference to the Black-Scholes prices for longer maturities. The differences plotted in the Figures 3.8–3.11 correspond to those in Figure 2.4 for the GH Lévy model. All simulation results are based on parameter estimates from the same Bayer return time series; this allows for comparability of the results in the Chapters 2 and 3.

Figure 3.8 shows the difference of IGOU prices minus Black-Scholes prices for different values of $\tau$. Recall, that a decreasing $\tau$ means a dependence structure with a longer range. Obviously, changes in $\tau$ affect the prices especially at-the-money with less than 60 days to expiration: A more slowly decreasing autocorrelation function leads to a steeper increase in the price difference.

Figure 3.9 explains the sensitivity of IGOU prices with respect to the kurtosis of the NIG distribution. A larger $\chi$, or equivalently a smaller $\zeta$, corresponds to a higher kurtosis. Obviously a higher kurtosis leads to an overall increase in the price differences.

In Figure 3.10 we have repeated the simulation results for the corresponding GH model and the supIGOU model. In the exponential Lévy model the price differences to Black-Scholes are large just before the expiration of the option whereas the difference increases in the IGOU model with time to maturity. The supIGOU prices are quite close to the prices in the IGOU model. This is also stressed in Figure 3.11 which summarizes all simulation results for the supIGOU and IGOU model with varying autocorrelation parameter $\tau$.

## Saddle-Point Approximation

In Section 3.8 we proposed the saddle-point approximation as an alternative to the time-consuming simulation of prices. Here we compute the saddle-point approximation to IGOU prices again for a simulated option data set.

The condition $z \in (-a_-, a_+)$ given in Proposition 3.90 for the existence of the cumulant generating function is sometimes violated for options with a very short time, i.e. one day, to
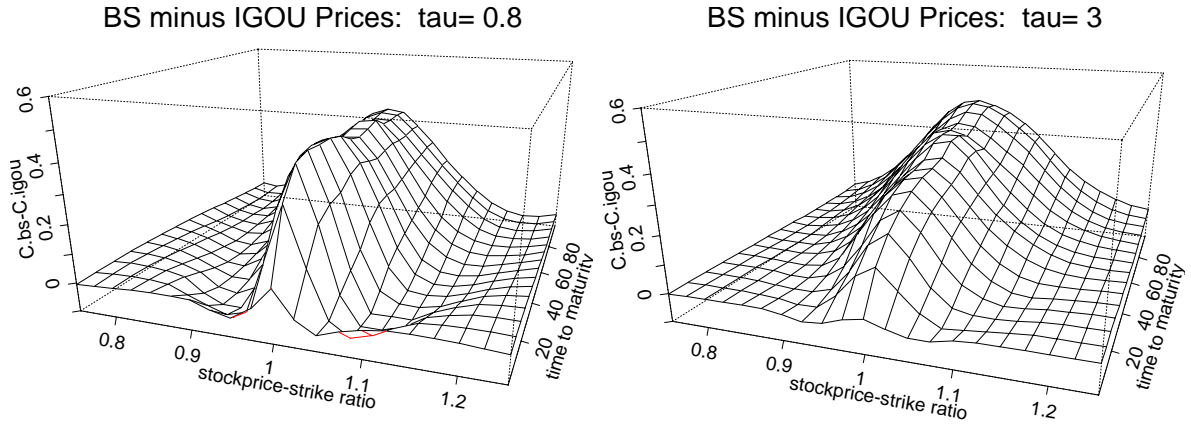
Figure 3.8: *Simulation of the difference of Black-Scholes minus IGOU prices. NIG distribution estimated for Bayer (see Table 1.14) and strike price K = 1000. $\tau = 1.143$ is estimated for Bayer.*
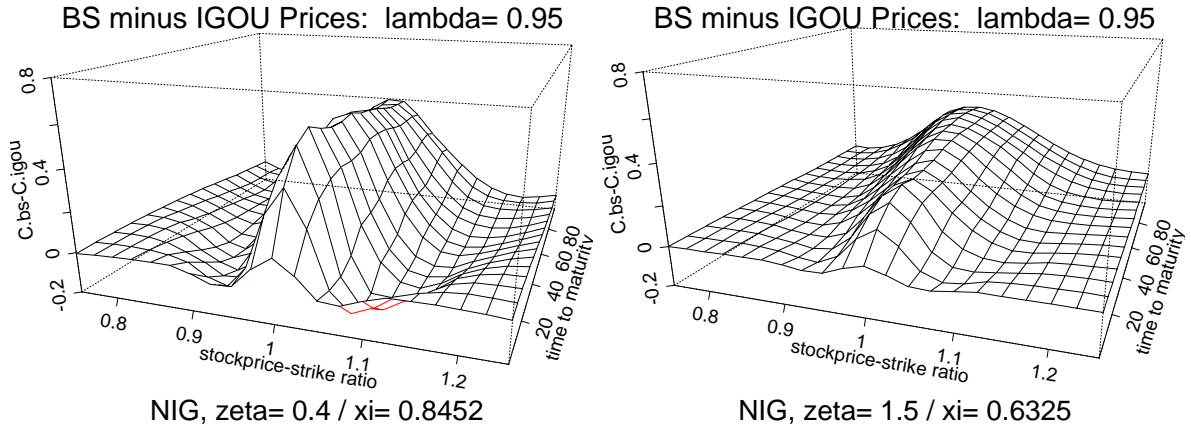


Figure 3.9: *Sensitivity of IGOU prices to the kurtosis of the NIG return distribution (strike price K=1000).*
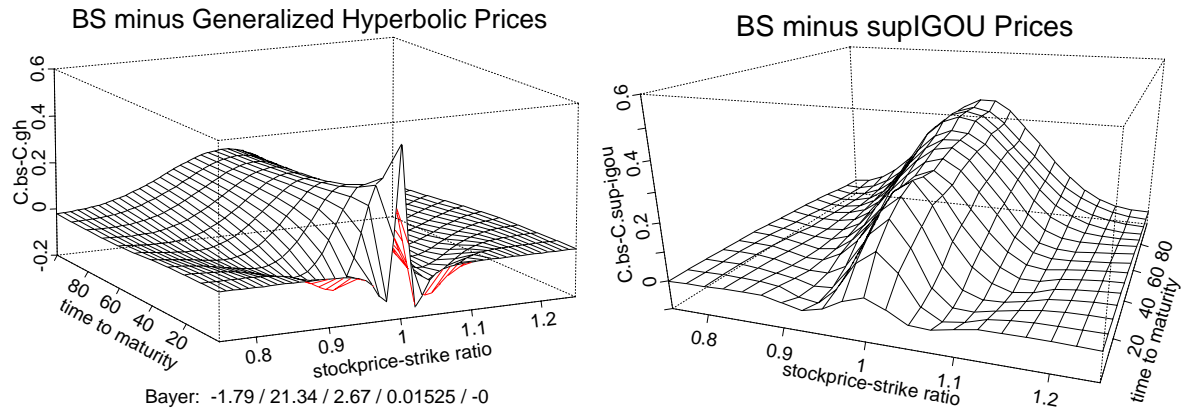


Figure 3.10: *Black-Scholes minus GH (left) and BS minus supIGOU prices (right). Estimates for Bayer stock returns: $\tau_1 = 0.0103$ / $\tau_2 = 1.7663$ / $m_1 = 0.1537$ / strike price K= 1000.*
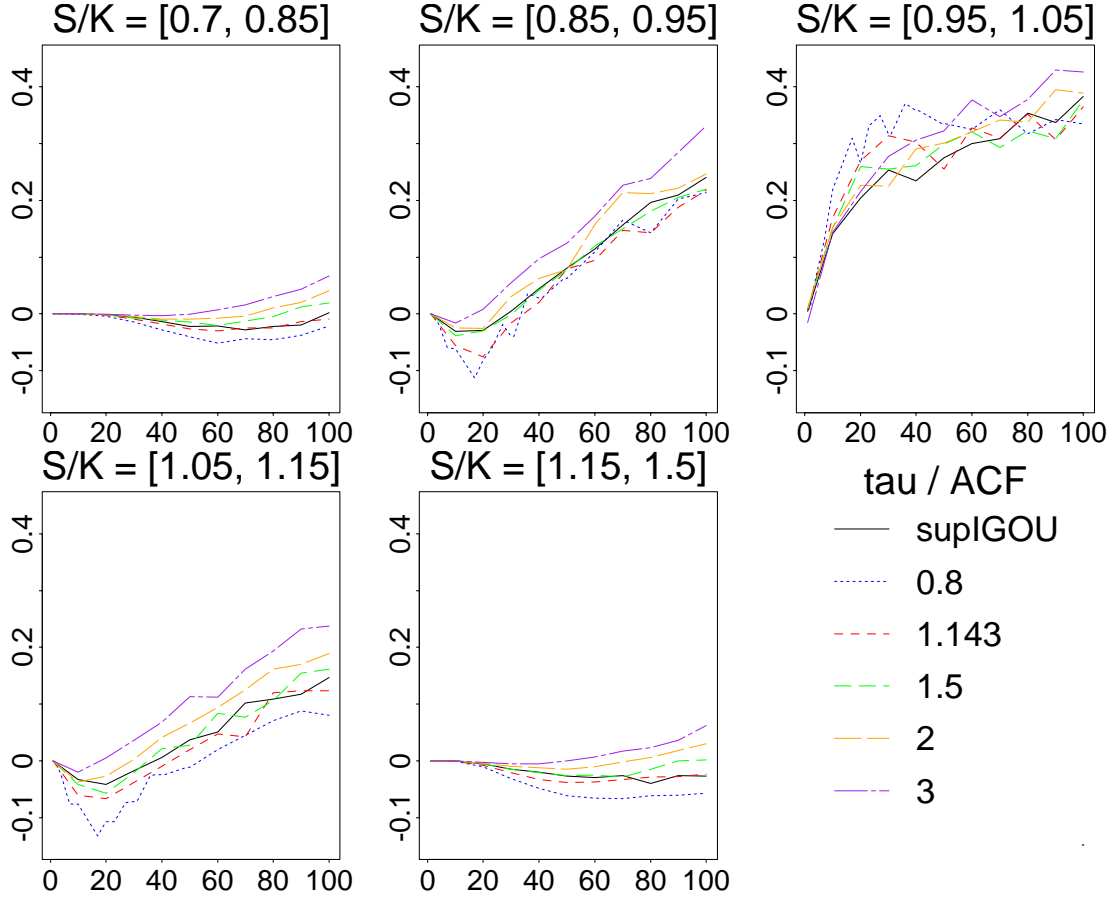
Figure 3.11: *Prices in different stockprice-strike regions: Black-Scholes minus IGOU prices, strike K=1000.*

maturity. We omit those options from the data set. From a practical point of view this is not of a deep concern, since we could—as a feasible alternative for short-term options—compute the prices via simulation. Moreover, the implementation of the saddle-point approximation is intricate because sometimes the numerical values are close to zero. One has to be careful that values which are close to zero are not set equal to zero.

On the other hand, the use of saddle-point approximations reduces the computational time to fractions of a second.

Figure 3.12 compares the IGOU prices based on the simulation of the integrated volatility and those computed via saddle-point approximation. The IGOU prices computed with both methods are close everywhere but for options in-the-money with a short time to maturity. The good results for the approximation in all but one region indicate that there are no large numerical errors in the other regions left. Only the difference of simulated and approximated prices in the problematic region has to be explained by more fundamental errors in one of the computation methods. The following remarks hint at possible sources of the error:

- The integrated volatilities are simulated for each time to maturity without respect to the stockprice-strike ration. Therefore, the error in terms of integrated volatility is the same for all options with the same time to maturity. However, because of the leverage
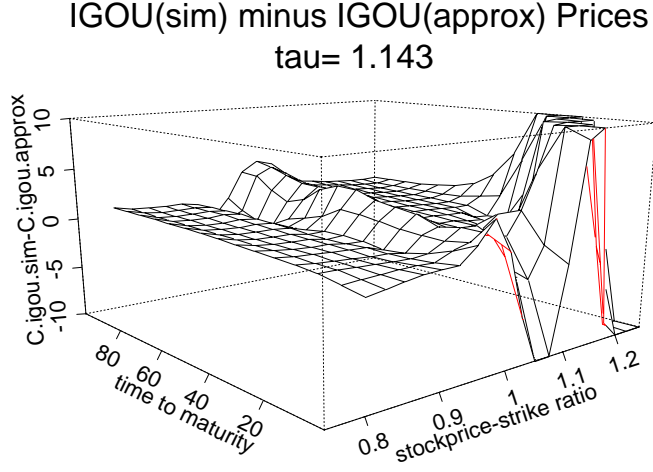
Figure 3.12: *Approximation error of the saddle-point approach to the computation of IGOU call prices (NIG distribution estimated from Bayer stock returns and strike price $K = 1000$).*

effect, the errors in terms of prices need not to be symmetric. The asymmetry of the errors in Figure 3.12 alone is no indication that the saddle-point results are wrong.

- The risk premium for options is largest for options at-the-money, whereas the price for call options far-out or far-in-the-money should be close to the intrinsic value ($S_0 - Ke^{-rT})^+$. This is the case for the prices computed by simulation. The approximated prices deviate substantially from the intrinsic value. This indicates that the saddle-point approximation of prices of in-the-money options close to maturity is not reliable enough.

From a practical point of view it could be useful to apply the saddle-point approximation for all options but those in-the-money with a short time to maturity.

The asymptotical properties of the saddle-point methods are valid for $n \to \infty$. In Section 1.14 we have applied this method for $n \geq 1$ corresponding to the number of trading days. In the IGOU model we apply it for $n = 1$. This may cause the approximation errors described above.

It would be interesting to investigate higher-order saddle-point approximations given in Lugannani and Rice (1980) and compare their numerical errors with the previous results.

## 3.12 Pricing Performance

In the preceding section we have shown that the derivative prices in IGOU type models may explain the smile effect and correct the often observed overpricing of options with long maturities in the Black-Scholes model. Following the simulated pricing comparisons in the last section, we are now evaluating the pricing performance of the model with real data. The test is based on intraday call option data sets for Bayer and the corresponding stock prices. See Chapter 2 for a more detailed description of the data set and the pricing performance of models based on GH Lévy motions.

94

Table 3.13: *Pricing errors with different volatility estimators: Mean and standard deviation of the difference of the model price minus quoted price, Bayer call options.*

| estimator | Black-Scholes | | NIG | | IGOU | | supIGOU | |
|---|---|---|---|---|---|---|---|---|
| | mean | st.dev. | mean | st.dev. | mean | st.dev. | mean | st.dev. |
| **Bayer calls, January 1993, 424 obs.** | | | | | | | | |
| Hist30 | −0.126 | 1.23 | −0.126 | 1.23 | −0.156 | 1.22 | −0.158 | 1.216 |
| Imp.median30 | 0.548 | 1.552 | 0.543 | 1.551 | 0.508 | 1.538 | 0.506 | 1.533 |
| **Bayer calls, February 1993, 662 obs.** | | | | | | | | |
| Hist30 | 0.606 | 1.779 | 0.599 | 1.779 | 0.564 | 1.764 | | |
| Imp.median30 | 0.715 | 1.817 | 0.706 | 1.817 | 0.67 | 1.802 | | |
| **Bayer calls, March 1993, 759 obs.** | | | | | | | | |
| Hist30 | 2.079 | 3.075 | 2.06 | 3.072 | 2.012 | 3.049 | | |
| Imp.median30 | 1.111 | 2.566 | 1.099 | 2.561 | 1.055 | 2.541 | | |

Our recipe to find a particular price is the the same as described in Section 2.1: We estimate parameters of the NIG distribution from daily data over a long-term period (1988–94). In the sequel we keep the scale- and location-invariant parameters $(\zeta, \varrho)$ and in particular the kurtosis of the NIG distribution fixed. This reflects the fact that the tail behaviour can only be estimated over a long period, e.g. several years. On the other hand we have to adapt to short-term volatility changes. We estimate the volatility in two ways: Hist30 is the standard deviation of the stock price 30 calendar days before the trade and Imp.median30 is the median of the 30 quoted Black-Scholes implied volatilities observed prior to a particular trade. We rescale the NIG distribution while keeping the shape parameters fixed, such that the variance of the NIG distribution is equal to the volatility estimate. See Section 2.1 for details of the rescaling procedure in the case of GH distributions. We use the mixing IG distribution of the NIG distribution as the marginal distribution of the IG Ornstein-Uhlenbeck type volatility process. Then we simulate the distribution $\lambda^P$ of the integrated volatility corresponding to this IGOU process with a long-term estimate (1988–94) of the autocorrelation parameter $\lambda$. Finally, we compute the price with Corrollary 3.80. Essentially the same procedure is used for supIGOU prices. We have not applied the saddle-point approximation method to investigate the pricing performance.

The results of the pricing experiment are given in Table 3.13 for three different monthly datasets of Bayer call options. The first two columns show the pricing errors and the standard deviation in the Black-Scholes model and under the assumption that log-prices follow a NIG Lévy motion. The risk-neutral measure is derived by the Esscher transform (see Section 1.11). For the IGOU model we often observe a reduction of pricing errors and of standard deviations. Although the differences are small, in all but one case the prices are corrected in the right direction.

Table 3.13 contains also results concerning the pricing performance of the supIGOU model in one particular month. From the simulation study in Section 3.11 (see for instance Figure 3.10) it is clear that we could not expect results which are substantially different to the results in the IGOU model. Obviously, the better modelling of the long-range dependence by the supIGOU model does not change the pricing behaviour markedly.

## 3.13 Statistical Martingale Measures

In this section we outline the *implied* approach to the estimation of parameters involved in the Ornstein-Uhlenbeck type process $\sigma^2$. For numerical reasons we compute the statistical martingale measure only for the Gamma OU model. Suppose, that the one-dimensional marginal distribution $\sigma_t^2 \sim \text{Gamma}(\lambda, \delta)$, with $\lambda > 1$ and $\delta > 0$. In this case we want to estimate $\lambda, \delta$ as well as the parameter $\tau$. A possible approach is as follows.

1. Given the market prices of $n$ derivatives $\widehat{C}_1$, $\widehat{C}_2$, ..., $\widehat{C}_n$, we use the IGOU model in order to compute theoretical prices $C_{\text{Gamma},1}(\lambda, \delta, \tau), \ldots, C_{\text{Gamma},n}(\lambda, \delta, \tau)$ of the same derivatives. If the quoted derivatives are European call options, we can use Corrollary 3.80 for pricing.

2. Then we can estimate the parameters $\lambda, \delta, \tau$ by minimizing pricing errors

$$\sum_{i=1}^{n} \left( C_{\text{Gamma},i}(\lambda, \delta, \tau) - \widehat{C}_i \right)^2. \tag{3.98}$$

In the Black-Scholes world this approach corresponds to the computation of implied volatilities, since the volatility is the only unknown parameter. See Eberlein, Keller, and Prause (1998, Chapter IX) for this approach in the case of the hyperbolic model.

Table 3.14: *Statistical martingale measures in the volatility model of the Gamma Ornstein-Uhlenbeck type computed for Bayer stock options.*

| Data set | obs. | $\gamma$ | $\lambda$ | $\tau$ | pricing error |
|----------|------|----------|-----------|--------|---------------|
| July 1–9, 1992 | 100 | 299.2942 | 1.999 | 14.525 | 19.132 |
| January 3–31, 1994 | 1240 | 153.298 | 1.473 | 0.73 | 1830.499 |

The *statistical martingale measures* (SMM) obtained for two data sets are given in Table 3.14. Since we have calculated the statistical martingale measures only for small data sets, we could not expect as stable results as in Section 2.6, where we used a two year data set. In Section 3.11 we have shown that the kurtosis of the NIG distribution and hence, the mixing IG distribution has a more procounced effect on the price than the autocorrelation parameter $\tau$. Therefore, we get quite unstable results for $\tau$.

The SMM Gamma distributions are shown in Figure 3.15. They are as distinct to the Black-Scholes implicit volatilities as the mixing IG distributions of NIG estimates from historical time series data (see also Figure 3.1).

Since historical and implicit volatilites are often quite different—in particular for the short time periods we used to estimate the statistical martingale measures—it is difficult to say something about the quality of the IGOU model based on these risk-neutral density estimates. (Note, that also the smile patterns shown in Section 2.4 are more erratic for monthly data sets. Since we wanted to show the general structure, we have not shown smiles for monthly data.)
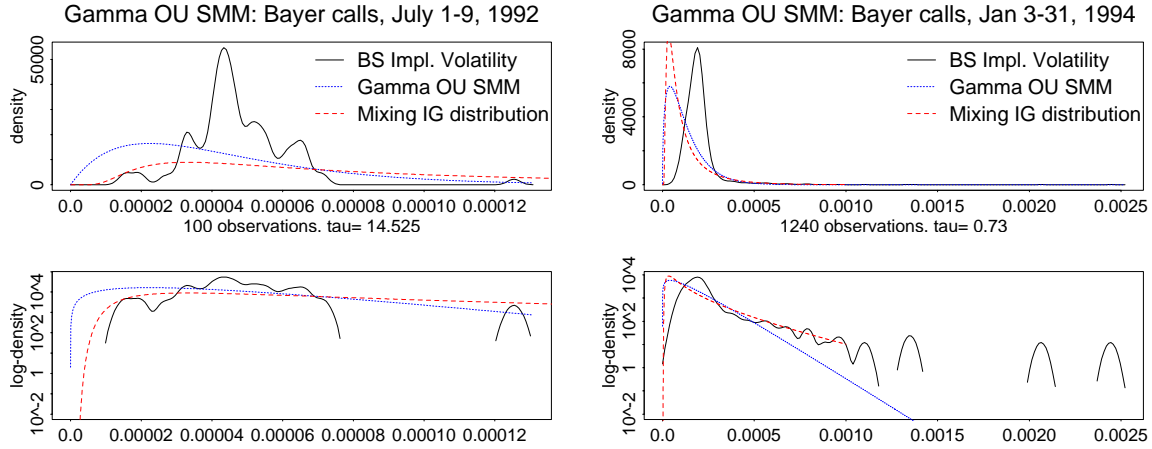
Figure 3.15: *Statistical martingale measures in the volatility model of the Gamma Ornstein-Uhlenbeck type computed for Bayer stock options.*

## 3.14 Volatility Models driven by Lévy Processes

We propose a refinement of the IGOU model by replacing the price process itself, above driven by a Brownian motion, by an exponential GH Lévy model

$$S_t = \exp(\sigma_t X_t) \tag{3.99}$$

$$d\sigma_t^2 = -\tau\sigma_t^2 dt + dZ_{\tau t}, \tag{3.100}$$

where $(X_t)_{0 \leq t \leq T}$ is a standard GH Lévy motion (i.e. with $\mathrm{Var} X_1 = 1$) and $(\sigma_t)_{0 \leq t \leq T}$ the IG Ornstein-Uhlenbeck type process described above. This GH–IGOU model combines a realistic modelling of the volatility structure by an IG Ornstein-Uhlenbeck type volatility process and the flexible modelling of the tails by a GH distribution.

A similar construction is proposed by Marinelli, Rachev, Roll, and Göppl (1999) for intraday prices of the Deutsche Bank stock: They consider a Gamma distribution for the description of the increments of the subordinated business time process (see also Marinelli, Rachev, and Roll (1999) for results concerning FX rates). Recall, that Gamma distributions are a special case of GIG distributions (see Table 1.1). Statistical tests hint at the presence of long-range dependence in their business time, which is simply the arrival time of price quotes in a particular market. Moreover, they reject a modelling of the business time process by processes with heavy tails. For the price process itself Marinelli, Rachev, Roll, and Göppl (1999) propose a stable process whereas we suggest a GH Lévy process. Using GH distributions should simplify the estimation problems markedly. We reviewed the approach of Marinelli, Rachev, Roll, and Göppl (1999) with the intention to show that the proposed new model does not come out of the blue.

Unfortunately, the marginal distribution of the log-prices does not belong to the class of GH laws. Therefore, we cannot obtain the IG distribution of the volatility process as a mixing distribution via formula (1.6). A reasonable procedure to estimate parameters is to estimate the volatility of observed prices locally, rescale the log-returns with the volatility (devolatize) and compute the GH likelihood estimate from the residuals.

The simulation of a price process is also straightforward: We can simulate the volatility process $\sigma^2$ and the NIG Lévy process $X$ independently. Here we simulate the volatiliy process
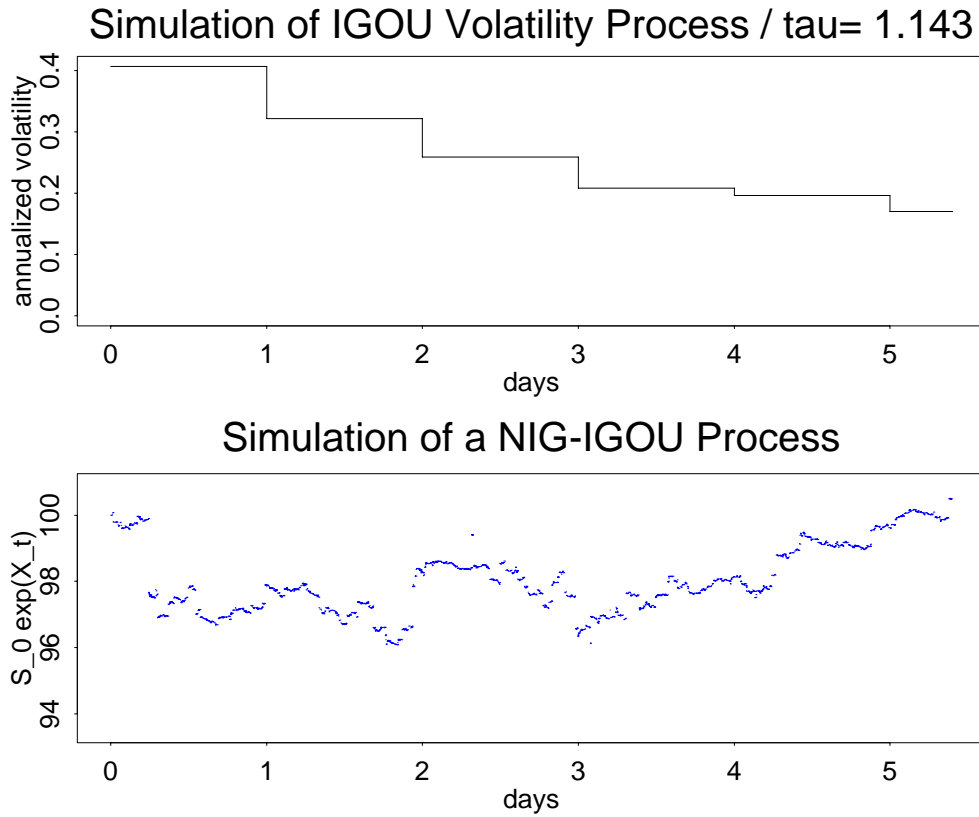
Figure 3.16: *Path of a simulated IG Ornstein-Uhlenbeck type process with parameters $\delta = 0.0116$, $\gamma = 72.3695$ and autocorrelation parameter $\tau = 1.143$ (top) and simulated geometric NIG–IGOU price process (bottom).*

as in Section 3.11 using the Rosinski (1991) expansion (with $\Delta = 1$ day); the standard NIG Lévy process is simulated pathwise by sampling a Poisson process and the jumps from the (truncated) Lévy measure independently. The path of the NIG Lévy motion used for Figure 3.16 was produced by Wiesendorfer Zahn (1999). This simulation approach to NIG Lévy processes was initially investigated by Rydberg (1997a).

In principle it is also possible to find an equivalent martingale measure with the same argumentation as in Section 3.6 and by the way to extend the Esscher approach from Chapter 1: Conditionally on the independent volatility process $\sigma^2$ an Esscher transform may be applied to find a risk-neutral density. This measure transformation is obviously equivalent on a finite time horizon. In analogy to Section 3.6 it should be possible to prove that if $X$ is a NIG Lévy motion the distribution of the log-price $\sigma_T X_T$ at maturity obeys a NIG law conditionally on $\sigma_T^{*T}$. If this holds, the calculation of "conditional" Esscher prices in the NIG–IGOU model is straightforward. However, there is no immediately obvious motivation for this *ad hoc* approach, for instance by hedging arguments.

# Chapter 4

# Multivariate Models
# in Finance

The challenge in the multivariate modelling of financial assets is that the models have to take the interaction of different assets into account. Statistically this is done by correlation matrices.

As a first step we give some insights into multivariate financial data: We use statistical methods in an explorative sense to describe the structure of e.g. the German stocks included in the DAX. In the sequel we also look at data sets with more complex structures. Although the methods in the first section are used in an explorative way, we will recline upon them for the construction and discussion of risk measures in Chapter 5. Then we will introduce multivariate GH distributions and propose an approach to estimate these distributions also for high-dimensional data efficiently. Finally we discuss the generation of random variates and the pricing of derivatives following the Esscher approach.

## 4.1  Correlation Structure

### Cluster Analysis

Cluster analysis provides the possibility to find structures in multivariate data without any previously given information: "A cluster should exhibit the properties of internal cohesion and external isolation" (Cormack 1971). A natural measure for the similarity is correlation. Therefore, we propose a distance adapted to returns of financial assets

$$ -\frac{1}{2}\big(\mathrm{Corr}(X_i, X_j) - 1\big)_{i,j \leq d}\, . \tag{4.1} $$

Here $(X_i)$ denote the returns resp. the absolute returns of stocks $i = 1, \ldots, d$. Note, that this metric needs a modification in the case of FX data because then the sign of the correlation has no meaning. The hierarchical clustering we apply here starts with a cluster for each object. We rely on the complete linkage approach, i.e. the distance between two clusters is defined as the largest dissimilarity between a member of cluster 1 and a member of cluster 2. Iteratively we fuse those two clusters that are closest (Seber 1984). From the resulting tree one obtains particular clusters by cutting the tree by a horizontal line. Though it is well-known that the hierarchical clustering is not a stable algorithm our approach gives an initial insight into the
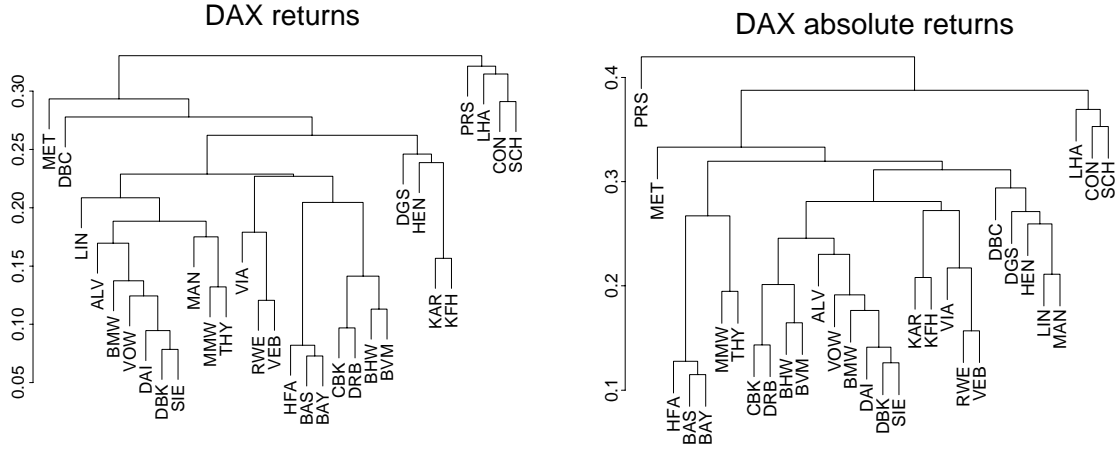
Figure 4.1: *Hierarchical clustering of returns and absolute returns.*

correlation pattern of the returns and the absolute returns. Note the striking similarity in the clustering observed for returns and absolute returns.

Looking at Figure 4.1, it is easy to find clusters with firms which belong to the same branch (e.g. Commerzbank/Dresdner Bank or BASF/Bayer). We also find—what is no surprise—high correlations of Deutsche Bank with Siemens, Daimler Benz. Cluster analysis allows to group the financial assets from an empirical point of view—in contrast to a classification based on the products of the firms.

## Principal Component Analysis

Principal component analysis (PCA) is a technique usually applied in the field of dimension reduction of multivariate data (see Seber (1984) for an introduction and Loretan (1997) for applications to the measurement of risk in financial markets). This approach needs no model assumptions in contrast to the similar factor analysis. Nevertheless, PCA is only a meaningful procedure if "axis symmetry" is given, i.e. that the joint distribution must be symmetric about its axes. The financial assets in our study are nearly symmetric. PCA is based on eigenvalue decomposition $T'ST = diag(\lambda_1, \ldots, \lambda_d)$, $\lambda_1 \geq \cdots \geq \lambda_d \geq 0$ of the sample dispersion matrix $S$. The the $j$-th column $y_j$ of $T'(x - \mu)$ is called the $j$-th principal component (PC). These principal components are uncorrelated and we have $\text{Var}(y_j) = \lambda_j$.

The principal component analysis, applied to multivariate financial data, reveals that there is one dominating principal component in the German stock market which explains 57% of the variance in the DAX (DAFOX 70%). Loretan (1997) found a similar structure for common stocks and FX rates—in contrast to short term interest rates which seem to be less correlated. As an example for a less homogeneous market we also carry out a principal component analysis for a subset of the NMZF data set (see Appendix A.1 for a description of the data sets). We examine those 20 assets which are quoted in Deutsche Mark.

In Figure 4.2 we show the obtained loadings, i.e. the weight of single variables in the principal components and explained proportions of variance for the successive principal components. For the German market we see that the first principal component is dominant. All 30 assets are included in the first component with a comparable and positive weight. In each of the other 29 components only one or two assets are important. In contrast to this we find more than one dominating component in the NMZF data set. We also observe that the
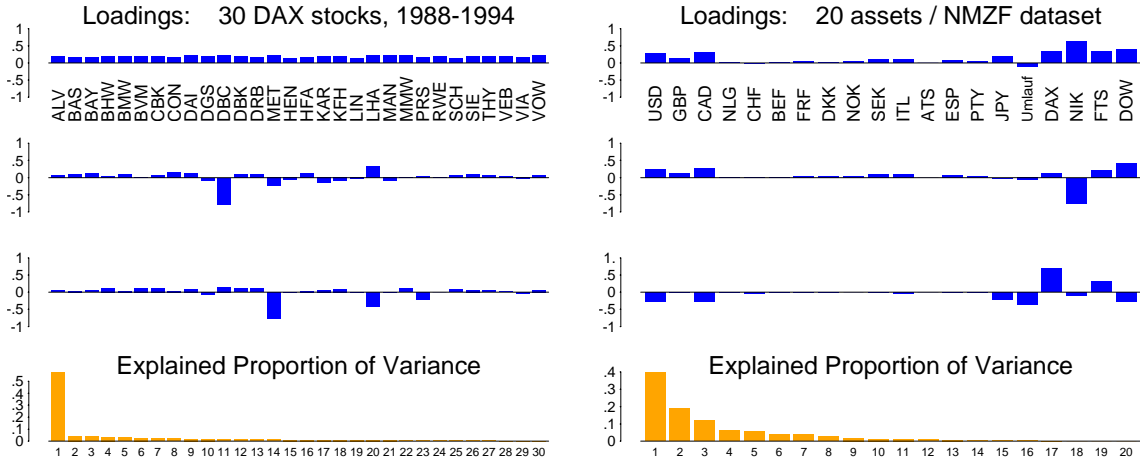
100

Figure 4.2: *Loadings of the first three principal components and explained proportion of variance resulting from the principal component analysis: 30 German stocks and 20 assets (15 FX rates, 1 interest rate ("Umlauf"/Umlaufrendite) and 4 indices). All values are quoted in Deutsche Mark resp. index points.*

interest rate has a negative weight and the magnitudes of the other weights are no longer comparable. For instance, the European FX rates have small loadings which is due to the reduced variability in the European foreign exchange market before the introduction of the Euro. The second PC explains opposite developments in the Japanese market to the US and European markets. And in all three plotted principal components the DAX and the interest rate "Umlaufrendite" have opposite weights. We do not want to interpret the economic meanings of single principal components here, but obviously we see more structure in the PCA of the NMZF data set. Loretan (1997) remarks also that it is often possible to find a useful, e.g. geographical, interpretation.

Another useful application of PCA in finance is the modelling of yield curves. In this case, the first three principal components are canonically interpreted as shift, twist and butterfly (Rodrigues 1997).

From the results of the PCA we conclude that the German stock market is very homogeneous in the following sense: The variance in the market is dominated by a single factor. The international market described the NMZF data set is less homogeneous. Therefore, it should be easier to diversify risk, i.e. using negative correlations to reduce the aggregated variance of the whole portfolio.

The motivation for the use of PCA in a financial context is dimension reduction. PCA is a useful tool for analyzing large portfolios which could consist of some hundreds of assets. It allows to find the source of high "values of risk" for a particular portfolio. Based on this analysis it is possible to develop a strategy to hedge risk. For example, an owner of a portfolio with German stocks included in the DAX may buy put options or invest in federal government bonds to reduce his exposure to market risk.

## The Epps Effect

In the discussion about correlations one has to add the following warning: In the case of a crash one observes that all equity prices are falling without exception. Therefore, the

"correlation" is close to one and no diversification by buying other equities in the same market is possible. This hints at limitations for correlation-based methods in measuring bankruptcy risk. From this point of view, the old method of banking regulators, who did not allow to use correlations in the portfolio to reduce the capital requirement, is not as ridiculous as it may sound. However, diversification of portfolios is also desirable in non-crash situations.

Epps (1979) already remarked that the correlation of intraday returns (10 min, 1 hour, ... ) is considerably lower than the correlation of daily returns. This complicates the construction of multivariate risk-measurement approaches which are applicable on different time scales.

## 4.2 Skewness and Kurtosis

Mardia (1970) proposed the following measures of multivariate skewness

$$b_{1,p} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left[ (X_i - \overline{X})' S^{-1} (X_j - \overline{X}) \right]^3 \tag{4.2}$$

and multivariate kurtosis

$$b_{2,p} = \frac{1}{n} \sum_{i=1}^{n} \left[ (X_i - \overline{X})' S^{-1} (X_i - \overline{X}) \right]^2, \tag{4.3}$$

where $\overline{X}$ denotes the sample mean vector and $S$ the sample covariance matrix. Mardia deduces the symptotic distributions of $b_{1,p}$ and $b_{2,p}$ under the assumption of normality; this allows for a test.

Table 4.3: *Multivariate kurtosis and skewness of daily returns. All corresponding p-values, calculated for the asymptotic distributions for $b_{1,p}$ and $b_{2,p}$ (Mardia 1970) are below 0.01%.*

|  |  | Empirical | Kurtosis Normal | var(kurt) | Skewness Empirical |
|---|---|---|---|---|---|
| DAX 1988-94 | 30 stocks | 1629.18 | 960 | 4.806 | 109.27 |
| NYSE 1990-96 | 4 subindexes | 58.49 | 24 | 0.11 | 0.538 |
| DAFOX 1974-95 | 12 indexes | 415.56 | 168 | 0.244 | 15.59 |

The results are given in Table 4.3. They indicate that there is skewness and kurtosis in multivariate financial data. The hypothesis of normality is rejected for the analyzed data.

We also have calculated the multivariate skewness and kurtosis for larger time lags than one trading day. The results are shown in Figure 4.4. Obviously, the kurtosis reduces for longer time lags. This is expected since, e.g., Lévy processes with jumps aggregate to Gaussianity. For larger time lags we observe first a slightly higher kurtosis which falls finally below the level of kurtosis of the normal distribution. This could be interpreted as the effect of stochastic volatility: Stochastic volatility models have instantaneous normal returns, after some time the stochastic volatility leads to an increase in the kurtosis, but in the long run most stochastic volatility models aggregate again to Gaussianity. Consequently, models with jumps and stochastic volatility (for instance the model proposed in Section 3.14) seem to be the appropriate models for multivariate financial data.
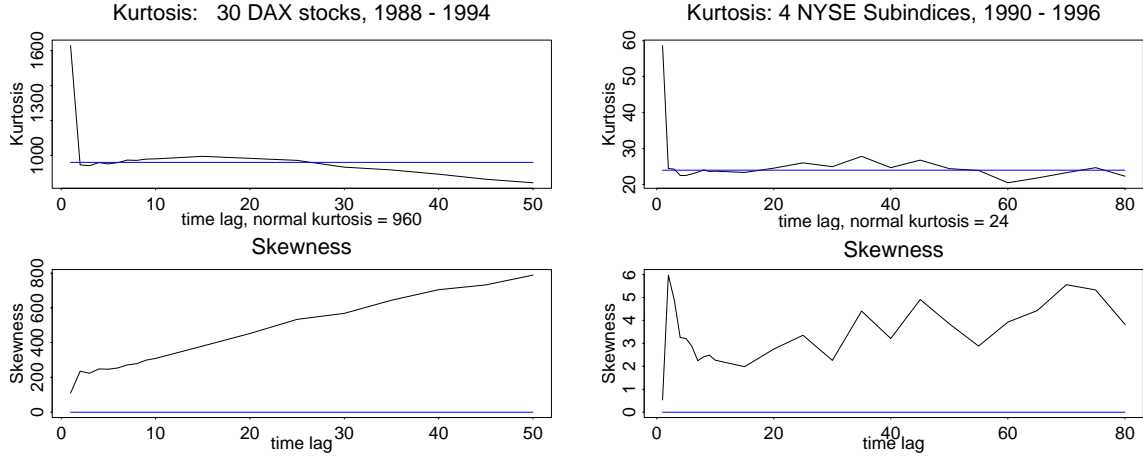
Figure 4.4: *Skewness and kurtosis for different time lags.*

## 4.3 Multivariate Generalized Hyperbolic Laws

In the preceeding sections we have described stylized features of multivariate financial data. Generalized hyperbolic distributions ($GH_d$) offer the possibility to construct models with the desired features; in particular, they allow to model the skewness ($\beta$ or $\Pi \in \mathbb{R}^d$) individually for every dimension. The kurtosis (described by $\lambda \in \mathbb{R}$ and $\zeta > 0$), i.e. the tail behaviour, is a result of the mixing distribution which is a univariate GIG law and, hence, similar in all dimensions. The fact that the multivariate GH distribution is also driven by a univariate mixing distribution—which might also be replaced by an appropriate volatility process— is sufficient for many modelling purposes in Finance, since the variance of all assets in a particular market usually changes together (see for instance the results of the PCA).

**Definition 4.4.** *The d-dimensional generalized hyperbolic distribution ($GH_d$) is defined for $x \in \mathbb{R}^d$ by the Lebesgue density*

$$\mathrm{gh}_d(x) = a_d \, \frac{\mathrm{K}_{\lambda-d/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)'\Delta^{-1}(x-\mu)}\,\right)}{\left(\alpha^{-1}\sqrt{\delta^2 + (x-\mu)'\Delta^{-1}(x-\mu)}\,\right)^{d/2-\lambda}} \, \exp\!\big(\beta'(x-\mu)\big),$$

$$a_d = a_d(\lambda,\alpha,\beta,\delta,\Delta) = \frac{\left(\sqrt{\alpha^2 - \beta'\Delta\beta}/\delta\right)^\lambda}{(2\pi)^{d/2}\,\mathrm{K}_\lambda\!\big(\delta\sqrt{\alpha^2 - \beta'\Delta\beta}\big)}$$

*These parameters have the following domain of variation[1]: $\lambda \in \mathbb{R}$, $\beta, \mu \in \mathbb{R}^d, \delta > 0$, $\alpha^2 > \beta'\Delta\beta$. The positive definite matrix $\Delta \in \mathbb{R}^{d\times d}$ has a determinant $|\Delta| = 1$. For $\lambda = (d+1)/2$ we obtain the multivariate* hyperbolic *and for $\lambda = -1/2$ the multivariate* normal inverse Gaussian *(NIG) distribution.*

Blæsild and Jensen (1981, p. 50) introduced a second parametrization $(\zeta, \Pi, \Sigma)$, where

$$\zeta = \delta\sqrt{\alpha^2 - \beta'\Delta\beta}, \quad \Pi = \beta\Delta^{1/2}\,(\alpha^2 - \beta'\Delta\beta)^{-1/2} \text{ and } \Sigma = \delta^2\Delta. \tag{4.5}$$

Corresponding to the univariate case (see Section 1.1) we can find simpler expressions in the cases $\lambda \in {}^1\!/_2\,\mathbb{Z}$. The symmetric hyperbolic and the symmetric normal inverse Gaussian

---

[1]We omit the limiting distributions obtained at the boundary of the parameter space; see for instance Blæsild and Jensen (1981). Generalized hyperbolic distributions are symmetric if and only if $\beta = \Pi = (0, \ldots, 0)'$

are defined in equation (4.17) and (4.18) below. Contour ellipses of GH distributions are computed in Blæsild (1981, pp. 254 ff).

To visualize the characteristics of these distributions and to compare the GH to the multivariate normal distribution we use the estimates of the following Section 4.4. Figure 4.5 (top left/right) shows the density and the log-density of the hyperbolic distribution estimated for Deutsche Bank/Volkswagen.

In the second row we compare the symmetric hyperbolic with the normal distribution. The left plot shows the difference of the densities. The right plot contains the logarithm of the absolute difference of the densities, which allows to compare the tail behaviour. In the center the hyperbolic distribution has more mass than the normal distribution, in the "middle" the normal distribution has higher densities and the tails of hyperbolic distributions are heavier again. The same pattern repeats for other GH distributions.

In the third row we compare the symmetric hyperbolic and the symmetric NIG distribution. In this case we have four different regions: The hyperbolic distribution has more mass in the center and the second circle; the NIG distribution has more mass in the first circle and the tails.

For the computation of moments and Esscher transforms, we need the multivariate moment generating function.

**Lemma 4.6.** *The moment generating function of the generalized hyperbolic distribution is given by*

$$
M(u) = e^{\mu u} \left( \frac{\alpha^2 - \beta'\Delta\beta}{\alpha^2 - (\beta + u)'\Delta(\beta + u)} \right)^{\lambda/2} \frac{\mathrm{K}_\lambda\big(\delta\sqrt{\alpha^2 - (\beta + u)'\Delta(\beta + u)}\,\big)}{\mathrm{K}_\lambda\big(\delta\sqrt{\alpha^2 - \beta'\Delta\beta}\,\big)},
$$

*where $\alpha^2 > (\beta + u)'\Delta(\beta + u))$.*

*Proof.* First, we assume without loss of any generality $\mu = 0$.

$$
\begin{aligned}
M(u) &= \int e^{u'x}\, \mathrm{gh}_d(x)d\lambda \\
&= a_d(\lambda, \delta, \kappa) \int \frac{\mathrm{K}_{\lambda-d/2}\big(\alpha\sqrt{\delta^2 + (x-\mu)'\Delta^{-1}(x-\mu)}\,\big)}{\big(\alpha^{-1}\sqrt{\delta^2 + (x-\mu)'\Delta^{-1}(x-\mu)}\,\big)^{d/2-\lambda}} \exp\big((\beta + u)'x\big) \\
&= \frac{a_d(\lambda, \delta, \kappa)}{a_d(\lambda, \delta, \sqrt{\alpha^2 - (\beta + u)'\Delta(\beta + u)})}
\end{aligned}
$$

Resubstituting $a_d$ and multiplying by $\exp(\mu u)$ yields the desired result. $\qquad\square$

**Lemma 4.7.** *In the case of a multivariate NIG distribution the moment-generating function given in Lemma 4.6 simplifies to*

$$
M_{\mathrm{NIG}}(u) = \exp\left[\mu u + \delta\left(\sqrt{\alpha^2 - \beta'\Delta\beta} - \sqrt{\alpha^2 - (\beta + u)'\Delta(\beta + u)}\right)\right].
$$

**Lemma 4.8.** *The characteristic function of the GH distribution is*

$$
\varphi(t) = \left( \frac{\alpha^2 - \beta'\Delta\beta}{\alpha^2 - \beta'\Delta\beta + \frac{1}{2}t'\Delta t - i\beta'\Delta t} \right)^{\lambda/2} \frac{\mathrm{K}_\lambda\big(\delta\sqrt{\alpha^2 - \beta'\Delta\beta + \frac{1}{2}t\Delta t - i\beta'\Delta t}\,\big)}{\mathrm{K}_\lambda(\delta\sqrt{\alpha^2 - \beta'\Delta\beta}\,)},
$$

*where $t \in \mathrm{I\!R}^d$.*
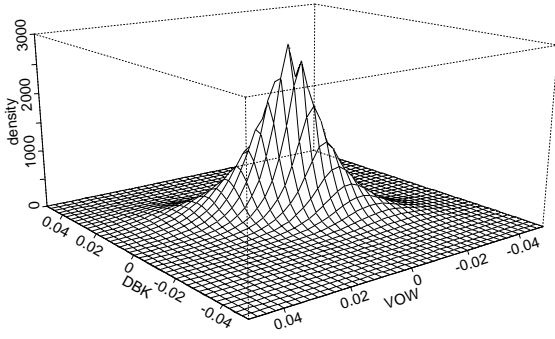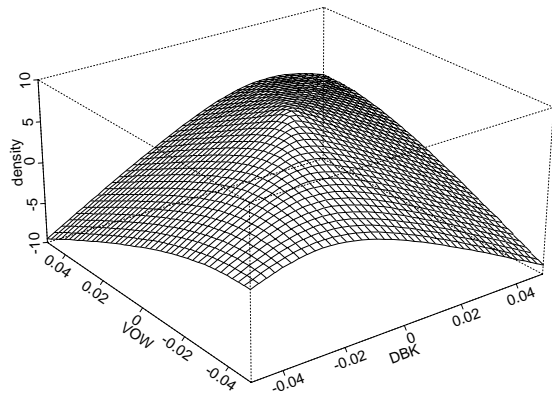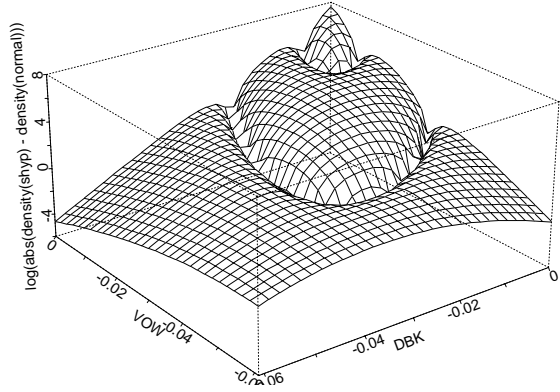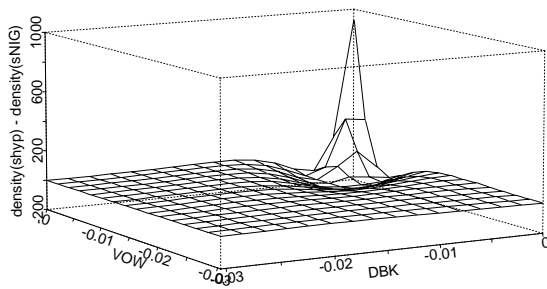
Figure 4.5: *Comparison of normal, hyperbolic and NIG densities: Deutsche Bank / Volkswagen.*

*Proof.* We use the characterization of the multivariate GH distributions as a mixture of a $d$-dimensional normal distribution with a GIG distribution (see Definition 1.7) and parameters $\chi = \delta^2$ and $\psi = \alpha^2 - \beta'\Delta\beta$. Following Barndorff-Nielsen and Halgreen (1977) the characteristic function of the GH distribution has the representation

$$\varphi(t) = e^{i\mu't}\zeta\left(\frac{1}{2}t'\Delta t - i\beta'\Delta t\right)$$

where $t \in \mathrm{I\!R}$ and $\zeta$ is the Laplace transform of the GIG distribution which is given in Lemma 1.39. Hence, we obtain
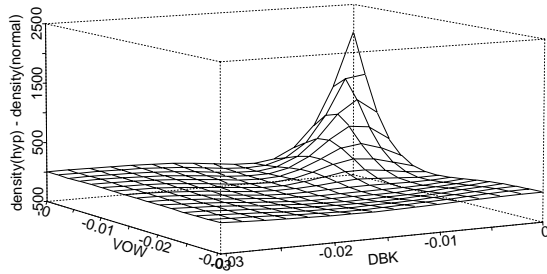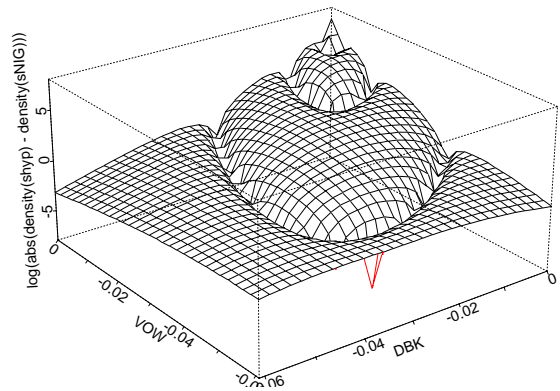
$$\varphi(t) = \left(\frac{\psi}{\psi + \frac{1}{2}t'\Delta t - i\beta'\Delta t}\right)^{\lambda/2} \frac{\mathrm{K}_\lambda\left(\sqrt{\chi(\psi + \frac{1}{2}t'\Delta t - i\beta'\Delta t)}\right)}{\mathrm{K}_\lambda(\sqrt{\chi\psi})};$$

replacing $\chi$ and $\psi$ yields the result. □

**Remark 4.9.** *Note, that Barndorff-Nielsen and Halgreen (1977) proved the infinite divisibility of multivariate GH distributions using the approach of the proof of Lemma 4.8: The GIG distribution is infinitely divisible and consequently with Feller (1966, p. 538) also the GH distribution is infinitely divisible. This allows the construction of $\mathrm{GH}_\mathrm{d}$ Lévy processes.*

**Remark 4.10.** *The multivariate characteristic function of the NIG distribution is easily computed using the equation $\mathrm{K}_{-1/2}(x) = \sqrt{\pi/2}\,x^{-1/2}\exp(-x)$. From the resulting representation it is clear that also the multivariate distributions $\{\mathrm{NIG}_d(\alpha, \beta, t\delta, t\mu, \Delta)\}$ for $t > 0$ are closed under convolution.*

We cite the following theorem which shows that generalized hyperbolic distributions are closed under margining, conditioning and regular affine transformations.

**Theorem 4.11.** *Suppose that $X$ is a $d$-dimensional variate distributed according to the generalized hyperbolic distribution $\mathrm{GH}_\mathrm{d}(\lambda, \alpha, \beta, \delta, \mu, \Delta)$. Let $(X_1, X_2)$ be a partitioning of $X$, let $r$ and $k$ denote the dimensions of $X_1$ and $X_2$, respectively, and let $(\beta_1, \beta_2)$ and $(\mu_1, \mu_2)$ be similar partitions of $\beta$ and $\mu$, let*

$$\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{pmatrix}$$

*be a partition of $\Delta$ such that $\Delta_{11}$ is a $r \times r$ matrix. Then one has the following:*

1. *The distribution of $X_1$ is the $r$-dimensional generalized hyperbolic distribution. $\mathrm{GH}_r(\lambda^*, \alpha^*, \beta^*, \delta^*, \mu^*, \Delta^*)$, where $\lambda = \lambda^*$, $\alpha^* = |\Delta_{11}|^{-1/(2r)}\big(\alpha^2 - \beta_2(\Delta_{22} - \Delta_{21}\Delta_{11}^{-1}\Delta_{12})\beta_2'\big)^{1/2}$, $\beta^* = \beta_1 + \beta_2\Delta_{21}\Delta_{11}^{-1}$, $\delta^* = \delta|\Delta_{11}|^{1/(2r)}$, $\mu^* = \mu_1$ and $\Delta^* = |\Delta|^{-1/r}\Delta_{11}$.*

2. *The conditional distribution of $X_2$ given $X_1 = x_1$ is the $k$-dimensional generalized hyperbolic distribution $\mathrm{GH}_k\big(\widetilde{\lambda}, \widetilde{\alpha}, \widetilde{\beta}, \widetilde{\delta}, \widetilde{\mu}, \widetilde{\Delta}\big)$, where $\widetilde{\lambda} = \lambda - r/2$, $\widetilde{\alpha} = \alpha|\Delta_{11}|^{1/(2k)}$, $\widetilde{\beta} = \beta_2$, $\widetilde{\delta} = |\Delta_{11}|^{-1/(2k)}\big(\delta^2 + (x_1 - \mu_1)\Delta_{11}^{-1}(x_1 - \mu_1)'\big)^{1/2}$, $\widetilde{\mu} = \mu_2 + (x_1 - \mu_1)\Delta_{11}^{-1}\Delta_{12}$ and $\widetilde{\Delta} = |\Delta_{11}|^{1/k}(\Delta_{22} - \Delta_{21}\Delta_{11}^{-1}\Delta_{12})$.*

3. *Let $Y = XA + B$ be a regular affine transformation of $X$ and let $||A||$ denote the absolute value of the determinant of $A$. The distribution of $Y$ is the $d$-dimensional generalized hyperbolic distribution $\mathrm{GH}_d(\lambda^+, \alpha^+, \delta^+, \mu^+, \Delta^+)$, where $\lambda^+ = \lambda$, $\alpha^+ = \alpha ||A||^{-1/d}$, $\beta^+ = \beta(A^{-1})'$, $\delta^+ = ||A||^{1/d}$, $\mu^+ = \mu A + B$ and $\Delta^+ = ||A||^{-2/d} A' \Delta A$.*

*Proof.* See Blæsild (1981, Theorem I). □

A consequence of this theorem is that we could find easily an invariant parametrization of the shape (Blæsild and Jensen 1981): The families of the $GH_d$ distribution with fixed parameters $(\lambda, \zeta, \Pi, \Sigma)$ defined in (4.5) are invariant under a transformations of location and scale.

We define the following functions to simplify the notation. See Appendix B for relevant properties of these functions.

$$R_\lambda(x) \quad = \quad \frac{\mathrm{K}_{\lambda+1}(x)}{\mathrm{K}_\lambda(x)} \tag{4.12}$$

$$S_\lambda(x) \quad = \quad \frac{\mathrm{K}_{\lambda+2}(x)\,\mathrm{K}_\lambda(x) - \mathrm{K}_{\lambda+1}^2(x)}{\mathrm{K}_\lambda^2(x)} \tag{4.13}$$

Then we obtain for the mean and variance of $X \sim GH_d$ (Blæsild and Jensen 1981)

$$\mathrm{E}\,X \quad = \quad \mu + \delta R_\lambda(\zeta)\Pi\Delta^{1/2} \tag{4.14}$$

$$\mathrm{Var}\,X \quad = \quad \delta^2 \left( \zeta^{-1} R_\lambda(\zeta)\Delta + S_\lambda(\zeta)\big(\Pi\Delta^{1/2}\big)'\big(\Pi\Delta^{1/2}\big) \right). \tag{4.15}$$

Note, that the expressions for the mean and the variance are much simpler in the symmetric case, i.e. for $\beta = \Pi = (0, \dots, 0)'$. Also for NIG distributions (4.14) and (4.15) simplify considerably, since $R_{-1/2}(x) = 1$ and $S_{-1/2}(x) = x^{-1}$ (see (B.33); the second identity follows immediately from (B.17)).

## 4.4 Estimation of Symmetric GH Distributions

For the estimation of hyperbolic distributions exists an algorithm, the `hyp` program developed by Blæsild and Sørensen (1992). This program allows to estimate skewed and noncentered hyperbolic distributions up to 3 dimensions in a reasonable time. Here we would like to estimate GH distributions for high-dimensional financial data also in an efficient way. Therefore, we propose to restrict the estimation to symmetric GH distributions. The assumed symmetry is reasonable for returns of financial assets. Symmetric GH distributions are also closed under conditioning, margining and regular affine transformation; this is a direct consequence of Theorem 4.11.

In the symmetric centered case GH distributions have the following Lebesgue density

$$\frac{(\alpha/\delta)^\lambda}{(2\pi)^{d/2}\,\mathrm{K}_\lambda(\delta\alpha)} \frac{\mathrm{K}_{\lambda-\frac{d}{2}}\big(\alpha\sqrt{\delta^2 + x'\Delta^{-1}x}\big)}{\big(\alpha^{-1}\sqrt{\delta^2 + x'\Delta^{-1}x}\big)^{\frac{d}{2}-\lambda}} \tag{4.16}$$

The $d$-dimensional symmetric centered NIG distribution ($\lambda = -1/2$) is characterized by the Lebesgue density (the skewed NIG density is given in Barndorff-Nielsen 1997)

$$2\delta \left(\frac{\alpha}{2\pi}\right)^{\frac{d+1}{2}} \exp(\delta\alpha) \frac{\mathrm{K}_{(d+1)/2}\big(\alpha\sqrt{\delta^2 + x'\Delta^{-1}x}\,\big)}{(\delta^2 + x'\Delta^{-1}x)^{(d+1)/4}}; \tag{4.17}$$

and the multivariate hyperbolic distribution ($\lambda = (d+1)/2$) has the following density

$$\frac{(\alpha/\delta)^{(d+1)/2}(2\pi)^{(1-d)/2}}{2\alpha\,\mathrm{K}_{\frac{d+1}{2}}(\alpha\delta)}\exp\left(-\alpha\sqrt{\delta^2 + x'\Delta^{-1}x}\right). \tag{4.18}$$

The first step of the estimation follows a "method of moments" approach: we estimate the sample mean $\widehat{\mu} \in \mathrm{I\!R}^d$ and the sample dispersion matrix $S$ using canonical estimators. In the symmetric case equation (4.14) yields that $\mu$ is identical with the mean. Therefore, we replace the observations $x_i \in \mathrm{I\!R}^d$ by $(x_i - \widehat{\mu})_i$ in the sequel. From (4.15) we get the following estimate for $\Delta$

$$\widehat{\Delta} = \frac{\zeta}{\delta^2 R_\lambda(\zeta)}\,S = S\,|S|^{-1/n}. \tag{4.19}$$

Consequently, the matrix $\widehat{\Delta}$ may be computed by norming the sample dispersion matrix such that $|\widehat{\Delta}| = 1$.

The second step is to compute $y_i = x_i'\widehat{\Delta}^{-1}x_i$ from observations $x_i \in \mathrm{I\!R}^d$, $1 \le i \le n$. Then the log-likelihood function has the following form

$$\begin{aligned}
L(x;\lambda,\alpha,\delta) &= n\left[\frac{d}{2}\log\frac{\alpha}{2\pi} - \lambda\log\delta - \log\mathrm{K}_\lambda(\delta\alpha)\right] \\
&\quad + \sum_{i=1}^{n}\log\mathrm{K}_{\lambda-\frac{d}{2}}\left(\alpha\sqrt{\delta^2 + y_i}\right) \\
&\quad + \frac{1}{2}\left(\lambda - \frac{d}{2}\right)\sum_{i=1}^{n}\log\left(\delta^2 + y_i\right).
\end{aligned} \tag{4.20}$$

As in the univariate case, the computation of log-likelihood functions simplifies for $\lambda \in \frac{1}{2}\,\mathbb{Z}$. In the case of NIG distributions, i.e. for $\lambda = -1/2$ the number of Bessel functions $\mathrm{K}_\lambda$ to compute is reduced by one. In the case of hyperbolic and hyperboloid distributions we only have to compute one Bessel function instead of $n+1$. Since the computation of Bessel functions is the most time-consuming part, in the latter cases the computation is substantially faster. In the sequel we give the likelihood function and their derivatives for the hyperbolic and NIG subclasses, but we omit the tedious details of the calculations.

The last step in the estimation procedure is to maximize this log-likelihood function. We use again the Powell algorithm implemented by Özkan (1997) in Splus. We implemented the estimaton algorithm successfully for hyperbolic and NIG distributions. In the case of an arbitrary $\lambda$ we get numerical problems due to very low values of the involved Bessel functions $\mathrm{K}_\lambda$. The GH density contains Bessel functions of order $\lambda - d/2$. Standard numerical algorithms are not stable enough for Bessel functions with very high or very low order $\lambda$.

## Hyperbolic distribution

The log-likelihood function for hyperbolic distributions is

$$\begin{aligned}
L_{\mathrm{hyp}}(x,\alpha,\delta) &= n\left[\frac{d+1}{2}\log\frac{\alpha}{\delta} + \frac{1-d}{2}\log(2\pi) - \log(2\alpha) - \log\mathrm{K}_{\frac{d+1}{2}}(\delta\alpha)\right] \\
&\quad - \alpha\sum_{i=1}^{n}\sqrt{\delta^2 + y_i}.
\end{aligned} \tag{4.21}$$

The derivatives are expressed in terms of the functions $R$ and $S$. See Appendix B for the necessary properties.

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}L_{\mathrm{hyp}} = n\left[-\alpha^{-1} + \delta R_{\frac{d+1}{2}}(\delta\alpha)\right] - \sum_{i=1}^{n}\sqrt{\delta^2 + y_i} \tag{4.22}$$

$$\frac{\mathrm{d}}{\mathrm{d}\delta}L_{\mathrm{hyp}} = n\left[-\frac{d+1}{\delta} + \alpha R_{\frac{d+1}{2}}(\delta\alpha)\right] - \alpha\delta\sum_{i=1}^{n}(\delta^2 + y_i)^{-1/2}. \tag{4.23}$$

The second derivatives are

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\frac{\mathrm{d}}{\mathrm{d}\alpha}L_{\mathrm{hyp}} = n\left[\alpha^{-2} + \frac{\delta}{\alpha}R_{\frac{d+1}{2}}(\delta\alpha) - \delta^2 S_{\frac{d+1}{2}}(\delta\alpha)\right]$$

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\frac{\mathrm{d}}{\mathrm{d}\alpha}L_{\mathrm{hyp}} = n\left[2R_{\frac{d+1}{2}}(\delta\alpha) - \delta\alpha S_{\frac{d+1}{2}}(\delta\alpha)\right] - \delta\sum_{i=1}^{n}(\delta^2 + y_i)^{-1/2}$$

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\frac{\mathrm{d}}{\mathrm{d}\delta}L_{\mathrm{hyp}} = n\left[\frac{d+1}{\delta^2} + \frac{\alpha}{\delta}R_{\frac{d+1}{2}}(\delta\alpha) - \alpha^2 S_{\frac{d+1}{2}}(\delta\alpha)\right]$$
$$- \alpha\sum_{i=1}^{n}\left[(\delta^2 + y_i)^{-1/2} - \frac{\delta^2}{(\delta^2 + y_i)^{3/2}}\right].$$

## NIG Distribution

For NIG distributions we obtain the following log-likelihood function

$$L_{\mathrm{nig}}(x, \alpha, \delta) = n\left[\log(2\delta) + \delta\alpha + \frac{d+1}{2}\log\frac{\alpha}{2\pi}\right] \tag{4.24}$$
$$+ \sum_{i=1}^{n}\left[\log \mathrm{K}_{\frac{d+1}{2}}\left(\alpha\sqrt{\delta^2 + y_i}\right) - \frac{d+1}{4}\log(\delta^2 + y_i)\right].$$

The first derivatives are

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}L_{\mathrm{nig}} = n\left[\delta + \frac{d+1}{\alpha}\right] - \sum_{i=1}^{n}\sqrt{\delta^2 + y_i}\, R_{\frac{d+1}{2}}\left(\alpha\sqrt{\delta^2 + y_i}\right) \tag{4.25}$$

$$\frac{\mathrm{d}}{\mathrm{d}\delta}L_{\mathrm{nig}} = n(\delta^{-1} + \alpha) - \sum_{i=1}^{n}\frac{\alpha\delta}{\sqrt{\delta^2 + y_i}}R_{\frac{d+1}{2}}\left(\alpha\sqrt{\delta^2 + y_i}\right) \tag{4.26}$$

and the second derivatives are

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\frac{\mathrm{d}}{\mathrm{d}\alpha}L_{\mathrm{nig}} = -\frac{n(d+1)}{\alpha^2}$$
$$- \sum_{i=1}^{n}(\delta^2 + y_i)\left[\frac{R_{\frac{d+1}{2}}\left(\alpha\sqrt{\delta^2 + y_i}\right)}{\alpha\sqrt{\delta^2 + y_i}} - S_{\frac{d+1}{2}}\left(\alpha\sqrt{\delta^2 + y_i}\right)\right]$$

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\frac{\mathrm{d}}{\mathrm{d}\alpha}L_{\mathrm{nig}} = n - \sum_{i=1}^{n}\left[\frac{2\delta}{\sqrt{\delta^2 + y_i}}R_{\frac{d+1}{2}}\left(\alpha\sqrt{\delta^2 + y_i}\right) - \alpha\delta S_{\frac{d+1}{2}}\left(\alpha\sqrt{\delta^2 + y_i}\right)\right]$$

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\frac{\mathrm{d}}{\mathrm{d}\delta}L_{\mathrm{nig}} = -\frac{n}{\delta^2} + \alpha^2\delta^2\sum_{i=1}^{n}(\delta^2 + y_i)^{-1}\, S_{\frac{d+1}{2}}\left(\alpha\sqrt{\delta^2 + y_i}\right)$$
$$- \alpha\sum_{i=1}^{n}R_{\frac{d+1}{2}}\left(\alpha\sqrt{\delta^2 + y_i}\right)\left((\delta^2 + y_i)^{-1/2} + 2\delta^2(\delta^2 + y_i)^{-3/2}\right).$$

## Tail-Emphasized Estimation

As an alternative approach to the estimation of GH distributions we propose to minimize the following distance

$$\sum_{\alpha \in A} \big(q_{\mathrm{emp}}(\alpha) - q_{\mathrm{gh}}(\alpha, \zeta, \Sigma)\big)^2, \qquad A = \{0.01, 0.05, 0.95, 0.99\}. \tag{4.27}$$

The motivation to minimize this particular distance is clearly given by the computation of risk measures; see Section 5.6 for more details.

## Estimation Results

For a price process $S_t \in \mathbb{R}^d$ we define returns $x_t \in \mathbb{R}^d$ by

$$x_t^{(i)} = \big[S_t^{(i)} - S_{t-\Delta t}^{(i)}\big]/S_t^{(i)} \approx \log S_t^{(i)} - \log S_{t-\Delta t}^{(i)}, \quad 1 \le i \le d, \tag{4.28}$$

which are approximated by the log-returns defined in equation (1.22). The reason for this definition of the returns is that the return of an investment described by a vector $h \in \mathbb{R}^d$ is then simply given by $h' x_t$.

Table 4.6 contains the results for the estimates for various data sets and three types of generalized hyperbolic distributions: symmetric hyperbolic, symmetric NIG and skewed hyperbolic. Skewed hyperbolic distribution with at most three dimensions are estimated with the `hyp` programm implemented by Blæsild and Sørensen (1992).

In general, the likelihood of the NIG distributions is higher and the distance (4.27) to the empirical distributions is smaller than for hyperbolic distributions. However, the estimation of hyperbolic distributions is much faster because of the already mentioned simpler structure of the distribution. Of course the likelihood of skewed hyperbolic distributions is higher than symmetric hyperbolic distributions. Nevertheless, it is below the likelihood of NIG distributions.

For the hyperbolic distribution the parameter $\xi$ is often estimated close to 1, which indicates the highest kurtosis possible for hyperbolic distributions. A comparison of the $\xi$-values for the maximum-likelihood and the tail-emphasized estimate reveals no clear picture. For only half of the datasets we obtain a higher tail-emphasized estimate for the kurtosis, described by the values of $\xi$.

The marginal density of a GH distributions is obtained by Theorem 4.11. Typically, we obtain the pattern shown in Figure 4.7 for the densities and log-densities: The marginal distributions of hyperbolic and NIG distributions are closer to the empirical distribution than normal distributions. In the center marginals of hyperbolic distributions are closer to the empirical distribution but in the tails marginals of NIG distributions provide a better fit.

Table 4.6: *ML and tail-emphasized estimates of hyperbolic and NIG distributions. The estimates of the location $\mu$, skewness $\beta$ and variance matrix $\Delta$ are omitted. All but the* `hyp` *estimate for the first data set are symmetric, i.e.* $\beta = (0, \ldots, 0)'$.

| | GH Parameters | | | | | Method | Log-Likel./ |
| | $\lambda$ | $\alpha$ | $\delta$ | $\zeta$ | $\xi$ | | $L^2$-distance |
|---|---|---|---|---|---|---|---|
| German stock, January 1, 1988 - May 24, 1994 / 1598 observations | | | | | | | |
| Daimler Benz / Deutsche Bank / Thyssen | | | | | | | |
| Sym. NIG | −0.5 | 96.634 | 0.01 | 1.010 | 0.705 | SGH | 15418.17 |
| Sym. HYP | 2 | 197.953 | $8.0e-16$ | $1.5e-13$ | 1 | SGH | 15404.57 |
| HYP | 2 | 199.042 | $2.7e-07$ | $6.0e-05$ | 0.999 | 'hyp'-Prog. | 15415.94 |
| Sym. NIG | −0.5 | 153.310 | 0.0167 | 2.569 | 0.529 | tail-emph. | 8.54e-05 |
| Sym. HYP | 2 | 219.608 | 0.010 | 2.207 | 0.558 | tail-emph. | 9.32e-05 |
| NYSE Indices, January 2, 1990 to November 11, 1996 / 1746 observations | | | | | | | |
| Industrial-Transport-Utility-Finance | | | | | | | |
| Sym. NIG | −0.5 | 251.345 | 0.007 | 1.844 | 0.592 | SGH | 26819.20 |
| Sym. HYP | 2.5 | 416.154 | $1.3e-16$ | $5.4e-14$ | 1 | SGH | 26800.84 |
| Sym. NIG | −0.5 | 222.999 | 0.006 | 1.485 | 0.634 | tail-emph. | 3.75e-05 |
| Sym. HYP | 2.5 | 409.1709 | $4.75e-10$ | $1.94e-07$ | 0.999 | tail-emph. | 4.59e-05 |
| NMZF-Returns: dowdem-ftsdem-daxdem-nikdem / 1781 observations | | | | | | | |
| Sym. NIG | −0.5 | 133.089 | 0.014 | 1.842 | 0.593 | SGH | 22831.64 |
| Sym. HYP | 2.5 | 220.572 | $1.3e-16$ | $2.9e-14$ | 1 | SGH | 22815.54 |
| Sym. NIG | −0.5 | 107.681 | 0.011 | 1.232 | 0.669 | tail-emph. | 4.15e-05 |
| Sym. HYP | 2.5 | 216.877 | $3.55e-09$ | $7.71e-07$ | 0.999 | tail-emph. | 7.16e-05 |
| NMZF-Returns: usddem-jpydem-gbpdem-chfdem/ 1781 observations | | | | | | | |
| Sym. NIG | −0.5 | 194.075 | 0.004 | 0.8255 | 0.740 | SGH | 28752.62 |
| Sym. HYP | 2.5 | 506.344 | $1.0e-15$ | $5.5e-13$ | 1 | SGH | 28735.53 |
| Sym. NIG | −0.5 | 222.45 | 0.0048 | 1.083 | 0.692 | tail-emph. | 7.26e-06 |
| Sym. HYP | 2.5 | 477.954 | $1.63e-09$ | $7.82e-07$ | 0.999 | tail-emph. | 1.55e-05 |
| 20-dim NMZF-Returns/ 1781 observations | | | | | | | |
| Sym. NIG | −0.5 | 375.886 | 0.003 | 1.085 | 0.692 | SGH | 163347.72 |
| Sym. HYP | 10.5 | 1780.619 | $2.1e-24$ | $3.773e-21$ | 1 | SGH | 162099.19 |
| Sym. NIG | −0.5 | 355.202 | 0.002798 | 0.994 | 0.708 | tail-emph. | 0.000188 |
| Sym. HYP | 10.5 | 1632.506 | $1.7e-09$ | $2.84e-06$ | 0.999 | tail-emph. | 0.000795 |
| 30-dim dataset, German stocks, January 1, 1988 - May 24, 1994 / 1598 observations | | | | | | | |
| Sym. NIG | −0.5 | 178.916 | 0.014 | 2.587 | 0.527 | SGH | 160554.84 |
| Sym. HYP | 15.5 | 641.615 | $8.0e-16$ | $5.1e-13$ | 1 | SGH | 159783.58 |
| Sym. NIG | −0.5 | 521.98 | 0.0426 | 22.255 | 0.207 | tail-emph. | 0.01483 |
| Sym. HYP | 15.5 | 720.34 | 0.0299 | 21.55 | 0.210 | tail-emph. | 0.01485 |
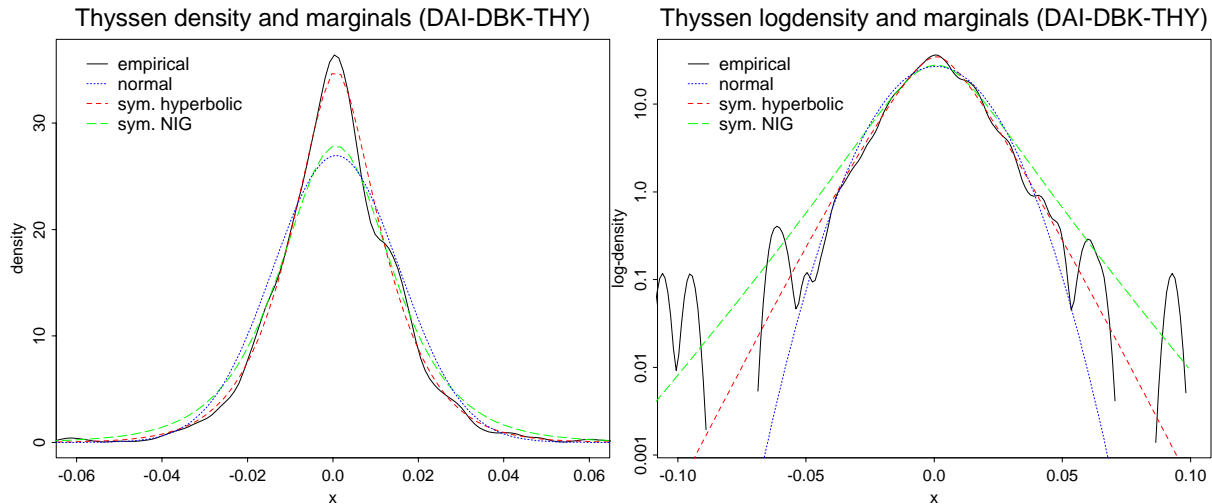
Figure 4.7: *Daimler Benz/Deutsche Bank/Thyssen. Marginals for Thyssen.*

## 4.5 Estimation of Skewed GH Distributions

In the previous section we have proposed a method to estimate high-dimensional *symmetric* generalized hyperbolic distributions, for instance from stock return data. A natural question is now, is it also possible to estimate high-dimensional *skewed* GH distributions efficiently? In principle, this is no problem, but the number of parameters increases too rapidly with the dimensions for an implementatin which is useful in practice. Financial models, in particular for market risk management, should allow a for fast calibration of the model and evaluation of the risk measures also for higher dimensions. Otherwise they are only of academic interest.

The `hyp` programm of Blæsild and Sørensen (1992) works very well for up to three dimensions, but for more dimensions the estimation of hyperbolic distributions seems not to be feasible in a reasonable time. Recall, that the hyperbolic distribution allows to construct the fastest estimation procedures of all GH distributions, since it contains only a Bessel function in the norming constant (see also Section 1.2 for remarks concerning the univariate estimation). For all other distributions from the GH family the estimation problems in higher dimensions are even worse, because the number of Bessel functions is greater or equal to the number of observations. Of course, for real-world financial applications three dimensions are not sufficient.

We have implemented an algorithm for the estimation of skewed GH distributions in higher dimensions—without encouraging results. The estimation is very slow and not stable. Moreover, numerical problems with the Bessel functions do occur frequently. Consequently, we decided to focus our attention to symmetric hyperbolic and symmetric NIG distributions. For both *symmetric* subclasses the algorithm is very reliable: For the backtesting studies over a six year time interval we could use the algorithm without numerical problems in a reasonable time (see Sections 5.4 and 5.6).

Nevertheless, we conclude with some remarks concerning the estimation of generalized hyperbolic distributions with skewness: First, it is useful to take the symmetric estimates as starting values. Usually financial log-returns are nearly symmetric, and therefore symmetric distributions are a very good approximation. Secondly, one has to keep in mind that the covariance matrix has to remain positive definite during the maximization of the log-likelihood

function. A solution to this problem is to decompose the matrix in two upper triangular matrices with positive diagonal elements. Their product is allways positive definite.

## 4.6 Generating Random Variates

In this section we construct random number generators for the multivariate generalized hyperbolic distribution. We examine four different ways to generate multivariate generalized hyperbolic random variates.

A straightforeward approach would be to sample from full multivariate distributions. Note, that it is not possible to find a transformation of the $GH_d$ distribution which yields independent marginal distributions. In the case of the normal distribution it is possible to find this transformation. Thus, the construction of a multivariate normal random generator reduces to sampling from univariate marginal distributions, which are fortunately also normal.

To compute $n$ random numbers $y_j \sim \mathrm{GH_d}(\lambda, \alpha, \beta, \delta, \mu, \Delta)$ we generate random variates $\{w_j : 1 \le j \le n\}$ with $w_j \sim \mathrm{GIG}\big(\lambda, \delta^2, \alpha^2 - h_0' \Delta h_0\big)$ and $\{x_i : 1 \le i \le dn\}$ with standard normal distribution. Then the $d$-dimensional GH random numbers are obtained as

$$y_j = \sqrt{w_j}\big(x_{d(j-1)+1}, \ldots, x_{dj}\big) D + \mu + w_j \beta \Delta, \tag{4.29}$$

where $D$ is the Cholesky decomposition of the matrix $\Delta$. Note that we have to compute the decomposition only once.

There are two possible ways to generate GIG random variates: First, one can evaluate the quantile function of GIG distributions (numerically) for iid uniform random variates. A second possibility is to use the algorithm proposed by Michael, Schucany, and Haas (1979), based on random variates with a normal distribution (see also Atkinson 1982).

An alternative is to draw random variates from univariate marginal distributions of conditional distributions. The first step in this approach is to sample from the univariate marginal distribution of the full $GH_d$ distribution (for the first dimension). Secondly, we have to calculate the distribution for the dimensions 2 to $d$ conditional on the first dimension; then we have to calculate the marginal distribution for the second dimension of the previously calculated conditional distribution. In the third step we would have to calculate the distribution for dimensions 3 to $d$ conditional on the first and second dimension and so on. Clearly this approach leads to a complicated computer program which has to be adapted to varying dimensions of multivariate GH distributions.

As a third possibility is to follow a Markov chain Monte Carlo (MCMC) approach which is well-known in Bayesian statistics especially for sampling from a posterior distribution. Two important examples are the Metropolis algorithm and the Gibbs sampler (Tierney 1994; Gilks 1994).

We propose to construct a Gibbs sampler for a given $d$-dimensional probability measure $\pi$ in the following way: For the target density $\pi(x), x \in \mathbb{R}^d$ we have to know the induced 1-dimensional conditional densities $\pi(x_i | x_{-1})$, where $x_{-i} = (\ldots, x_{i-1}, x_{i+1}, \ldots)$. We begin with an arbitrary starting value $x^0 \in \mathbb{R}^d$ and obtain a sequence of realisations $x^1, x^2, \ldots$ with $x^t \in \mathbb{R}^d$ by successive drawings of random variates for each $x_i^t$ from the conditional probability densities $\pi(\cdot | x_1^t, x_{i-1}^t, x_{i+1}^{t-1}, x_d^{t-1})$. We denote with $K_G \colon \mathbb{R}^d \times \mathbb{R}^d \to [0,1]$ the

transition kernel of the Markov chain $X = (X^1, X^2, \dots)$ given by

$$K_G(x, y) = \begin{cases} \prod_{\ell=1}^{d} \pi(y_\ell | x_j, y_i, i < j) & \text{if } \int \pi(y_1, \dots, y_i, x_{i+1}, \dots, x_k) dy_i > 0 \\ 0 & \text{otherwise} \end{cases}$$

For the proof of the convergence we need the following notation: We define $D = \{x \in E : \pi(x) > 0\}$ and a function $h \colon \mathbb{R}^d \to \mathbb{R}_+$ as lower semicontinuous at 0 iff for all $x$ with $h(x) > 0$ there exists an open neighbourhood $B$ of $x$ and $\varepsilon > 0$ such that for all $y \in B$ holds $h(y) \geq \varepsilon > 0$. The $n$-step Gibbs sampler is then defined as

$$K_G^n(x_n, x_0) = \int \cdots \int K_G(x_n, x_{n-1}) K_G(dx_{n-1}, x_n) \cdots K_G(dx_1, x_0)$$

**Proposition 4.30.** *The $n$-step Gibbs sampler $K_G^n$ constructed by the 1-dimensional conditional density $\pi(x_i | x_{-1})$ given in Theorem 4.11 for a multivariate generalized hyperbolic density converges weakly against the GH target density $\pi$.*

*Proof.* GH distributions have a continuous Lebesgue density which is clearly a stronger property than being lower semicontinuous.

From Roberts and Smith (1994, Theorems 1 and 2) we know that if $\pi$ has a $d$-dimensional Lebesgue density, $\pi$ is lower semicontinuous at 0, $\int \pi(x) dx_i$ is locally bounded for $i = 1, \dots, d$, and D is connected then $\forall x \in D$ follows that $|K_G^n - \pi| \to 0$ as $n \to \infty$.

The result follows since $D = \mathbb{R}^d$ and $\int \pi(x) dx_i$ is locally bounded in the case of GH distributions. $\qquad\square$

The aim of this section is to provide a recipe to draw a sequence of independent random generates from multivariate GH distributions. Therefore, we define $P_n^l$ as the measure on $\mathbb{R}^{d\ell}$ generated by a finite number of Gibbs samplers $K_G^n$

$$P_n^\ell(B) = \int \cdots \int_{(x_n', x_{2n}', \dots, x_{\ell n}') \in B} K_G^n(dx_{\ell n}, x_{(\ell-1)n}) \dots K_G^n(dx_{2n}, x_n) K_G^n(dx_n, x_0).$$

**Proposition 4.31.** *The measure $P_n^\ell$ converges weakly against the product of the stationary distributions $\pi$ of the $n$-step Gibbs sampler $K_G^n$*

$$P_n^\ell \xrightarrow{\ n\ } \bigotimes_{i=1}^{\ell} \pi$$

*Proof.* We define $x^{(n)} = (x_n', x_{2n}', \dots, x_{\ell n}')$. Regard the characteristic function $\varphi_n^\ell$ of the measure $P_n^\ell$, which is defined for $t \in \mathbb{R}^{d\ell}$. We apply Proposition 4.30 to get the limiting characteristic function

$$\begin{aligned} \varphi_n^\ell(t) \quad &= \quad \int e^{it'x^{(n)}} dP_n^\ell(x^{(n)}) \\ &= \quad \int \cdots \int e^{it'x^{(n)}} K_G^n(dx_{\ell n}, x_{(\ell-1)n}) \cdots K_G^n(dx_{2n}, x_n) K_G^n(dx_n, x_0) \\ &\xrightarrow{\ n\ } \quad \int \cdots \int e^{it'x^{(\infty)}} \pi(dx) \cdots \pi(dx) \end{aligned}$$

where $x \in \mathbb{R}^d$ and $x^{(\infty)} \in \mathbb{R}^{d\ell}$. The proposition follows with Cuppens (1975, Theorem 2.6.9). $\qquad\square$

The previous theorem provides the theoretical possiblity to construct a random generator for multivariate GH distributions. For the practical implementation we recommend to follow three steps:

1. Find the affine transformation $XA + B$ for a random variate $X \sim GH_d$ in Theorem 4.11 to simplify the parameters of the GH distribution. After the simplification it is sufficient to implement the random generator for a GH distribution with parameters $(\lambda, \alpha^+, \beta^+, \delta^+, 0, I_d)$.

2. The Gibbs sampler is than constructed using the 1-dimensional conditional distributions (see Theorem 4.11) $\pi(x_i|x_{-i}) = \overline{GH}_{d-1}$ with parameters $\bar{\lambda} = \lambda - \frac{1}{2}$, $\bar{\alpha} = \alpha^+$, $\bar{\beta} = \beta^+_{-1}$, $\bar{\delta} = \{(\delta^+)^2 + x'_{-i}x_{-i}\}^{1/2}$, $\bar{\mu} = 0$ and $\bar{\Delta} = 1$.

3. Finally, we have to transform the obtained $d$-dimensional random variates $(x_i)_{i\geq1}$ to $(x_iA^{-1} - B)_{i\geq1}$.

Note, that the $\pi(x_i|x_{-i})$ distribution is a 1-dimensional centered GH distribution and that the observations $x_{-i}$ enter the conditional distribution via the parameter $\delta$ as an inner product. Since $\delta$ is the scaling parameter, the previous observations affect especially the variance and not the "shape" of the distribution $\overline{GH}$.

It is worth to mention that $\pi(x_i|x_{-i})$ is hyperbolic if and only if $\pi(x)$ is hyperbolic. This reduces the necessary computational efford. In the case of NIG distributions we get a $GH_{d-1}$ distribution with $\bar{\lambda} = -1$.

## 4.7 Pricing of Derivatives

**Multivariate Esscher Transformations**

We start the usual filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{0\leq t\leq T}, P)$, where $\mathcal{F}$ is generated by the $d$-dimensional process $(S_t)_{0\leq t\leq T}$ to be introduced below and satisfies the usual conditions. Prices of $d$ non-dividend paying assets are given by $(S_t^j)_{0\leq t\leq T, 1\leq j\leq d}$. As usual, we define the log-prices as $X_t^j = \ln(S_t^j/S_0^j)$. Let $X_t = (X_t^1, \ldots, X_t^d)'$ be the vector of log-prices and $r$ the constant interest rate. We model $(X_t)_{0\leq t\leq T}$ by a $d$-dimensional (homogeneous) Lévy process which is of course characterized by the infinitely divisible distribution of $X_1$.

In this section we desribe the multivariate Esscher transforms introduced by Gerber and Shiu (1994) for option pricing purposes. In contrast to the original approach, we look only at a finite time horizon $0 \leq t \leq T$ to assure the equivalence of the martingale measure.

We define the cumulative density function $F(x,t) = P(X_t \leq x)$ for $x \in \mathbb{R}^d$ and the moment-generating function $M(z,t) = \mathrm{E}[\exp(z'X_t)]$ for $0 \leq t \leq T$ and $z \in \mathbb{R}^d$. Assume, that $(X_t)_{0\leq t\leq T}$ is a stochastic process with iid increments. Consequently $M(z,t) = [M(z)]^t, 0 \leq t \leq T$, where $M(z)$ is defined as the moment generating function of the infinitely divisible distribution $\mathcal{L}(X_1)$. We also assume that $(X_t)_{0\leq t\leq T}$ has a continuous density given by

$$f(x,t) = \frac{d^n}{dx_1 \ldots dx_n}F(x,t), \quad 0 \leq t \leq T. \qquad (4.32)$$

The multivariate Esscher transform for parameter $h \in \mathbb{R}^d$ is then defined by

$$f(x,t,h) = \frac{\exp(h'x)}{M(h,t)}f(x,t), \quad \text{where } x, h \in \mathbb{R}^d. \qquad (4.33)$$

The moment generating function under the Esscher transformed measure $P^h$ is

$$M(z,t,h) = \int \exp(z'X_t)dP^* = \int \frac{\exp(z'X_t + h'X_t)}{M(h,t)}\, dP$$

$$= \frac{1}{M(h,t)} \int \exp(z'X_t + h'X_t)\, dP = \frac{M(z+h,t)}{M(h,t)}. \tag{4.34}$$

**Proposition 4.35.** *Let $X$ be a multivariate Lévy process and assume that its moment generating function exists for $z \in U$ with $U$ open and $0 \in U \subset \mathrm{I\!R}^d$. Define*

$$\varrho_t^h = \frac{dQ_{X_t}^h}{dP_{X_t}} = \exp\big(h'x - t \log M(h)\big), \quad \text{where} \quad |h| < C. \tag{4.36}$$

*Then $Z_t^h = \varrho_t^h \circ X_t$ is a positive P-martingale and $Z^h$ defines for every $T < \infty$ a new measure $Q_T^h$ on $\mathcal{F}_T^X$ (the Esscher transform) equivalent to $P_T$ given by $dQ_T^h = Z_T^h dP_T$ such that the process $(X_t)_{0 \leq t \leq T}$ is a Lévy process under $Q_T$.*

*Proof.* The Propositions 19 and 20 in Keller (1997) hold after mild modifications also in the case of multivariate Lévy processes. $\quad\square$

We have seen that Esscher transforms describe an equivalent change of measure and the transformed process is again a Lévy process. Consequently, we have $M(z,t,h) = [M(z,1,h)]^t$.

Keller (1997) discusses parametric changes of measures by changing triplets. Most of the results hold also in the multivariate case.[2] Especially his Lemma 21 holds, which means that the Esscher transform of a Lévy process with triplet $(b,0,G)$ can be described by changing triplets. Under the integrability condition $\int_{|x| \geq 1} \exp(h'x)dG < \infty$ the Esscher transform is given by the Lévy triplet

$$\Big(b' = b + \int h(y-1)dG, c' = 0, G' = yG\Big), \tag{4.37}$$

where $y$ is defined as $y(x) = \exp(h'x)$ for $x, h \in \mathrm{I\!R}^d$. The triplet (4.37) characterizes the measure $Q$ uniquely. Incidentally, Esscher transforms of triplets and of marginal distributions $(X_t)_{0 \leq t \leq T}$ are essentially given by the same function.

Now we have to find the risk-neutral Esscher transform given by $h_0 \in \mathrm{I\!R}^d$, which makes the process $(e^{-rt}S_t)_{0 \leq t \leq T}$ to a martingale. We denote this measure by $Q$ and the corresponding expectation by $\mathrm{E}^Q$. Following Delbaen and Schachermayer (1994), we can use this equivalent martingale measure to compute prices for derivatives. Therefore, we have to find the parameter $h_0$ such that

$$S_0^j = \mathrm{E}^Q\big[e^{-rt}S_t^j\big] \quad \forall 0 \leq t \leq T,\ 1 \leq j \leq d. \tag{4.38}$$

Define $1_j = (0,\ldots,1,\ldots,0)$ where the 1 is on the $j$th position. Then we have that equation (4.38) is equivalent to

$$e^{rt} = \mathrm{E}^Q\big[S_t^j/S_0^j\big] = \mathrm{E}^Q\big[\exp(X_t^j)\big] = M(1_j, t, h_0)$$
$$= [M(1_j, 1, h)]^t$$
$$\Longleftrightarrow \quad r = \log M(1_j, 1, h_0)$$
$$\Longleftrightarrow \quad r = \log M(1_j + h_0) - \log M(h_0) \tag{4.39}$$

for all $0 \leq t \leq T$ and $1 \leq j \leq d$.

---

[2]We do not repeat the proofs in the multivariate case, because only minor changes are necessary.

The following theorem allows the computation of more complicated options, e.g. the value of an option to exchange one asset for another at the end of the period (Gerber and Shiu 1994, p. 123–126).

**Theorem 4.40 (Gerber and Shiu 1994, pp. 121–122).** *For a measurable function* $g$ : $\mathbb{R}^d \to \mathbb{R}$ *and* $0 \leq t \leq T$ *we have*

$$\mathrm{E}^{h_0}\big[e^{-rt}S_t^j g(S_t^1, \ldots, S_j^d)\big] = S_0^j \, \mathrm{E}^{h_0+1_j}\big[g(S_t^1, \ldots, S_j^d)\big],$$

*where* $\mathrm{E}^h$ *denotes expectation with respect to the Esscher transformed measure.*

## Esscher Transforms of Symmetric NIG Lévy Motions

Of course it is possible to apply the pricing approach described above to arbitrary multivariate GH Lévy motions, but we restrict ourselves to symmetric centered normal inverse Gaussian (NIG) Lévy motions, since we want to propose a model which is computationally feasible. See Prause (1999b) for a summary of the theoretical properties and empirical results for the subclass of symmetric NIG distributions. This subclass enjoyes an increased numerical and analytical tractability; hence, it can be used in real-time implementations to measure market risk and for the pricing of derivatives.

The moment generating function for symmetric centered NIG distributions is given by

$$M(u) = \exp\big[\delta\big(\alpha - \sqrt{\alpha^2 - u'\Delta u}\,\big)\big], \tag{4.41}$$

where $\alpha^2 > u'\Delta u$. We have to solve the $d$ equations given in (4.39). In this case they are equivalent to

$$r = \delta\Big(\sqrt{\alpha^2 - h_0{}'\Delta h_0} - \sqrt{\alpha^2 - (h_0 + 1_j)'\Delta(h_0 + 1_j)}\,\Big), \quad 1 \leq j \leq d. \tag{4.42}$$

Solving these equations leads to the risk-neutral Esscher parameter $h_0 \in \mathbb{R}^d$. The risk-neutral Esscher transform of $\mathcal{L}(X_1)$ is again a NIG distribution with Lebesgue density

$$\mathrm{nig}(x; \alpha, 0, \delta, 0, \Delta, h_0) = \frac{\exp(h_0{}'x)\,\mathrm{nig}(x)}{M(h_0)}. \tag{4.43}$$

NIG distributions are closed under convolution, consequently the risk-neutral distribution of $X_t$ is furthermore $\mathrm{NIG}(\alpha, h_0, \delta t, 0, \Delta)$ for $0 \leq t \leq T$.

Since we obtain a multivariate NIG distribution as a risk-neutral probability measure at the end of the period $[0, T]$, we can easily evaluate arbitrary terminal payoff functions. For payoff functions which do not allow analytic expressions we can also apply a Monte-Carlo approach to evaluate the discounted expectation under the risk-neutral measure. The procedure based on the mixing representation of GH distributions is described in Section 4.6.

## Skewed GH Lévy Motions

In the last paragraph we restricted ourselves to the special case of an asset pricing model driven by symmetric NIG Lévy motions. This model is very pleasant for the estimation of densities and the pricing of derivatives. However, if we look at the most general model, we can

state, that Esscher transforms of GH distributions are again GH distributions. The following computation shows, that we basically obtain a transformation of the skewness parameter.

$$
\begin{aligned}
\mathrm{gh}(\lambda, \alpha, \beta, \delta, \mu, \Delta; h_0)(x) &= \frac{\exp(h_0'x)}{M_{GH}(h_0)} \mathrm{gh}(\lambda, \alpha, \beta, \delta, \mu, \Delta)(x) \\
&= \frac{\exp(h_0'(x-\mu))}{M_{GH}(h_0)\exp(-h_0'\mu)} \mathrm{gh}(\lambda, \alpha, \beta, \delta, \mu, \Delta)(x) \\
&= \frac{\exp((h_0+\beta)'(x-\mu))}{M_{GH}(h_0)\exp(-h_0'\mu)} \mathrm{gh}(\lambda, \alpha, 0, \delta, \mu, \Delta)(x) \\
&= \mathrm{gh}(\lambda, \alpha, \beta + h_0, \delta, \mu, \Delta)(x)
\end{aligned}
$$

## 4.8 Basket Options

For a vector $w \in \mathbb{R}^d$ we consider a "basket" of assets $S_t^\dagger = \sum_{i=1}^d w_j S_t^j / S_0^j$ for $0 \le t \le T$. The weights $w$ describe the quantity of single assets in a portfolio with value process $S^\dagger$. We approximate this process by $S_t^\dagger \approx \exp(A_t)$ where $A_t$ is defined as

$$
A_t = \sum_{j=1}^n w_j X_t^j = w'X_t, \quad 0 \le t \le T. \tag{4.44}
$$

A "basket option" is now given by the terminal payoff function $(S^\dagger - K)_+$ for a strike price $K$. Consequently, we have to evaluate $e^{-rT} \mathrm{E}^*[(S^\dagger - K)_+]$. For a "real world" NIG Lévy motion with $\mathcal{L}(X_t) = \mathrm{NIG}(\alpha, 0, \delta, 0, \Delta)$ we obtain as a risk-neutral Esscher transform the Lévy motion with $X_t \sim \mathrm{NIG}(\alpha, h_0, t\delta, 0, \Delta)$, $0 \le t \le T$. For the distribution of $A_T$ under $Q_T$ we obtain a univariate $\mathrm{NIG}(\alpha|h'\Delta h|^{-1/2}, 0, T\delta|h'\Delta h|^{1/2}, 0)$ distribution.

One drawback is, that we use the approximation $S_t^\dagger \approx \exp(A_t)$ and not the exact value $S_t^\dagger = \sum_{j=1}^d \exp(X_t^j)$. However, using the mentioned Monte-Carlo approach, it is easy to calculate the price of a basket option without an approximation error (cf. Beißer (1999) for methods to reduce the approximation error in the calculation of prices for basket options).

One can also use the fact that NIG distributions are closed under marginalisation and regular affine transformation (this follows from Blæsild 1981, Theorem 1). This allows the efficient calculation of option prices based on these mappings. In contrast to this, conditional distributions are no longer normal inverse Gaussian. Only in the one-dimensional case we recommend the use of Fast Fourier Transforms to calculate densities of convolutions for GH distributions.

# Chapter 5

# Market Risk Controlling

What are the reasons for the increasing interest in risk measurement? Risk measures are used in banks and investment firms with different objectives. Banking regulators as well as the management want to reduce the probability of bankruptcy. Therefore, they set limits to the exposure to market risk relative to the capital of the firm. Internally, risk measures provide a possibility for the senior management to allocate risk:

> Setting limits in terms of risk helps business managers to allocate risk to those areas which they feel offer the most potential, or in which their firms' expertise is greatest. This motivates managers of multiple risk activities to favor risk reducing diversification strategies.[1]

J.P. Morgan and Reuters (1996) illustrate the allocation of risk to the subdivisions of a firm in their Chart 3.1.

On the one hand risk measures are a necessary simplification of a complex risk structure. Scalar Figures describing the exposure to market risk like VaR estimates satisfy this requirement. Moreover, they are easy to compare with the capital of the firm. On the other hand the methodology should allow a precise understanding what are governing factors leading to high potential losses. A methodology should also hint at possibilities to reduce risk, i.e. to diversify the portfolio or to insure the portfolio by buying or selling appropriate derivatives.

## 5.1  Concepts of Risk Measurement

The standard methodology in the finance industry to measure risk of a random variate $X$ is Value-at-Risk (VaR), i.e. the potential loss given a level of probability $\alpha \in (0, 1)$,

$$P[X < -\text{VaR}_\alpha] = \alpha.$$

Is VaR the adequate measure to meet the requirements mentioned above? Quantile-based methods like Value-at-Risk have the disadvantage, that they do not answer the question "How bad is bad?" The magnitude of losses greater than the chosen level of probability remains disregarded. However, not only the current risk-based capital rules for the market risk exposure of US commercial banks, effective as of January 1, 1998, are explicitly based on VaR estimates. See Basel Committee on Banking Supervision (1996) for the general rules

---

[1] J.P. Morgan and Reuters (1996, p. 33).

and Lopez (1999) for a description of the current evaluation methods and calculation of the capital requirement used by US regulatory authorities. From a supervisor's point of view, it is dangerous that institutions or even individual traders may establish a portfolio with a small measured risk to reduce the capital requirement—which leads to an inacceptable risk exposure e.g. below the 1%-quantile.[2] For the allocation of risk in a firm—in contrast to the prevention of bankruptcy—VaR is maybe more appropriate.

Stress testing offers a partial solution of this problem. This approach is focussing on extreme scenarios by simulating historical or possible future crash scenarios. From a regulator's point of view, evaluation methods for risk measurement should be comparable across institutions and applicable to all types of financial assets. These requirements are easier to fulfill with VaR-type approaches than with the construction of scenarios. On the other hand standardized stress scenarios are often not specific enough to identify all risks that pose a thread to a particular institution.

The Basel Committee on Banking Supervision (1995, Chapter V) demands that banks that use internal models for meeting market risk capital requirements must have a rigorous and comprehensive stress testing program that covers a range of factors that can create extraordinary losses in trading portfolios.

To avoid bankruptcy one has to forecast the distribution of the maximum expected loss. From this point of view regulators should use other risk measures than VaR. A better incorporation of extreme events especially in view of nonlinear portfolios is desirable. One concept (which is also easy to implement and to understand) is the *shortfall* defined as

$$\text{Shortfall}_{\alpha,t} = -E\big[X \big| X < q(\alpha)\big], \tag{5.1}$$

where $q : [0,1] \to \mathbb{R}$ is the corresponding quantile function. We assume that $E|X|$ exists. Note, that the shortfall goes clearly beyond the concept of VaR because it takes the extreme negative returns into account. The log-density of the empirical distribution in Figure 5.2 shows that they often occur at a level of probability below 5% or even below 1%.

## Coherent Risk Measures

We would like to mention the axiomatic concept of *coherent risk measures* developed by Artzner, Delbaen, Eber, and Heath (1999). Their article starts with a discussion of necessary properties of risk measures and reviews current approaches to risk management.

**Definition 5.2.** $\varrho : \mathbb{R}^d \to \mathbb{R}$ *is a* coherent risk measure *if and only if for all real random variates $X, Y$ and $n, t \in \mathbb{R}$ the following holds*

1. $\varrho(X + Y) < \varrho(X) + \varrho(Y)$, *Subadditivity,*

2. $\varrho(tX) = t\varrho(X)$, *Homogeneity,*

3. $\varrho(X) \geq \varrho(Y) \quad \forall X \leq Y$, *Monotonicity,*

4. $\varrho(X + rn) = \varrho(X) - n$, *Risk-Free Condition,*

---

[2]The Basel Committee on Banking Supervision (1999) is concerned that currently used models to cope with *Credit Risk* are not sophisticated enough. This allows the construction of portfolios which "exploit divergences between true economic risk and risk measured." Models for market risk are more advanced in comparison. However, the described shortcomings of the VaR approach do not inhibit the existence portfolios with a high risk which occurs only below a given level of probability.

*where r is the risk-free interest rate. A position X is defined as* acceptable *if and only if* $\varrho(X) \leq 0$.

Note that VaR is no coherent risk measure since it is not subadditive, whereas the shortfall is one (Artzner, Delbaen, Eber, and Heath 1997; see also Embrechts, Klüppelberg, and Mikosch 1997, Section 6.2.2, and Matthes and Schröder 1998 for an application of shortfall measures in practice).

The US Securities and Exchange Commision (SEC) rules, as applied by the National Association of Securities Dealers (NASD), are based on the construction of finitely many scenarios and belongs to the class of coherent risk measures (Artzner, Delbaen, Eber, and Heath 1997). In principle, this is the most challenging way to construct a risk measure, since it is necessary to understand all possible sources of risk. Artzner, Delbaen, Eber, and Heath (1997) prefer the generalized scenarios, because "it requires thinking before calculating, which can only improve risk management." However, the construction of generalized scenarios is not straightforward and dependent on the individual market structure and the modelled financial assets.

## Procedures to Compute Value-at-Risk

To compute Value-at-Risk, there are different ways how to get from an observation of a price vector $(S_t^1, \ldots, S_t^d)'$ to estimates of (scalar) risk measures. At first, one may fix the number of assets or the amount of money initially invested into a particular asset. Fixing the number of assets allows a reconstruction of the price path of the portfolio, which is denoted $\bar{S}_t^h$ in Figure 5.1. This approach is *portrayed* on the right side of the figure and, for instance, used in the construction of stock indices. Changes in the price of a particular stock alter also its weight measured in currency units for instance in the DAX.

On the contrary, one may keep the investment in each particular asset fixed (left side of Figure 5.1). In this case it is of course possible to reconstruct univariate return distributions $\mathcal{L}(\bar{X})$ of the portfolio value (center part of the Figure). The estimation of multivariate return distributions allows a simple computation of risk measures for a given portfolio denoted by $h \in \mathrm{I\!R}^d$. This approach is widely used in practice. For instance, J.P. Morgan and Reuters (1996) provide volatilities and correlations of most globally traded financial assets, i.e. they sell distribution estimates to investors and risk managers. With this information institutions can compute risk estimates for their portfolios without the need to handle and update large databases.

A small error occurs in the approximation of the portfolio return by a scalar product of log-returns. This approximation error can be removed by a Monte-Carlo simulation, which is of course time consuming. This approximation error is also of some consideration in the computation of basket option prices (see Section 4.8 and Beißer 1999).

The optimal choice of the risk-management method depends on some preferences, e.g. the simulation of $\check{S}$ gives more exact results but it is in particular for large data sets very computer-intensive. The reconstruction of the portfolio (right side of Figure 5.1) makes it in every case necessary to have the full data set available to adapt to changes in the portfolio composition. On the contrary, the updating procedure of the IGARCH approach leads to very small memory requirements (see Section 5.3 for the recursion formulas for the variance and covariance).

In the sequel we have considered the investment of a fixed amount of money, i.e. one currency unit, in each asset. Experiments indicate that the results in the sequel do not depend on the choice of the portfolio $h = (1, \ldots, 1)'$.
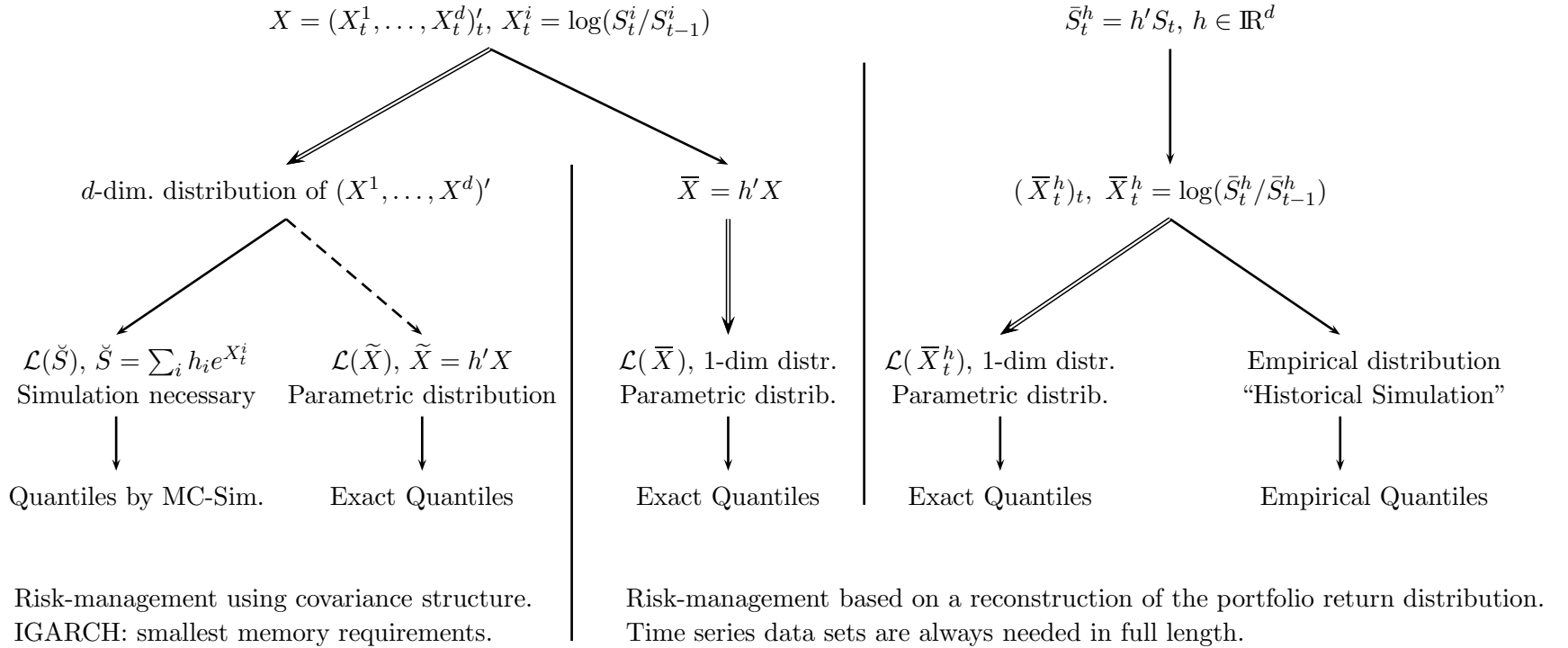
Figure 5.1: *Approaches to multi-dimensional risk-management.*

**Fixed Investments**
measured in currency units

$S = (S_t^1, \ldots, S_t^d)_t'$

**Variable Weights**
(number of assets fixed)

$X = (X_t^1, \ldots, X_t^d)_t',\ X_t^i = \log(S_t^i / S_{t-1}^i)$

$\bar{S}_t^h = h'S_t,\ h \in \mathbb{R}^d$

$d$-dim. distribution of $(X^1, \ldots, X^d)'$

$\overline{X} = h'X$

$(\overline{X}_t^h)_t,\ \overline{X}_t^h = \log(\bar{S}_t^h / \bar{S}_{t-1}^h)$

$\mathcal{L}(\check{S}),\ \check{S} = \sum_i h_i e^{X_t^i}$
Simulation necessary

$\mathcal{L}(\widetilde{X}),\ \widetilde{X} = h'X$
Parametric distribution

$\mathcal{L}(\overline{X})$, 1-dim distr.
Parametric distrib.

$\mathcal{L}(\overline{X}_t^h)$, 1-dim. distr.
Parametric distrib.

Empirical distribution
"Historical Simulation"

Quantiles by MC-Sim.

Exact Quantiles

Exact Quantiles

Exact Quantiles

Empirical Quantiles

Risk-management using covariance structure.
IGARCH: smallest memory requirements.

Risk-management based on a reconstruction of the portfolio return distribution.
Time series data sets are always needed in full length.

Remarks: $\Longrightarrow$ denotes estimation, e.g. following a maximum-likelihood approach.

$--\rightarrow$ 1-dim distribution obtained by matrix manipulation of the covariance matrix (Lemma 5.3, Theorem 5.4).

This scheme is common for generalized hyperbolic and normal distributions.

**Holding Period**

Usually capital requirements based on risk measures are calculated for holding periods of ten trading days (Basel Committee on Banking Supervision 1995). Due to the fact that trading takes place with different time horizons, the holding period should correspond to the trading frequency in a particular market. For instance, traders on global FX markets often close their positions every evening. Consequently. risk measures evaluated for overnight portfolios cannot capture the risk inherent in the day trading on FX markets. An often underestimated aspect is the liquidity of markets in the time of a crash. The time period necessary to close a position on a particular market is an indicator for the appropriate choice of the holding period to use for risk measurement.

## 5.2 Statistical Realization

In the last section we have discussed what are appropriate risk measures. In the sequel we describe standard methods to actually compute these measures and propose more realistic approaches based on generalized hyperbolic distributions.

A standard method to compute VaR is to simulate the return of the portfolio by the preceding, e.g. 250, observed returns and to use the quantile of this distribution. The Basel Committee calls this approach historical simulation. However, it is well-known that the estimation risk is higher for empirical quantiles than for quantiles of parametric distributions (Huisman, Koedijk, and Pownall 1998; see also Section 1.6). A proposed sophistication of this approach is to rescale the observed data set with up-to-date volatilities (Hull and White 1998).

A second simulation technique proposed to forecast VaR is Monte Carlo. It is computationally intensive, in particular, for large portfolios. Therefore, it is less widely used to measure the exposure to market risk (Basel Committee on Banking Supervision 1995, Paragraph I.7). We have not examined this method in the backtesting studies because distinct differences to the variance-covariance approach are only obtained for nonlinear portfolios (Bühler, Korn, and Schmidt 1998) and it should be adjusted to an individual risk-measurement problem. Hence, it offers too much variability for a valid statistical comparison. See Section 5.7 for a description of a full valuation approach based on GH distributions.

The variance-covariance methodology is based on the multivariate normal distribution and it is very popular since it is easy to calculate the distribution of an affine transformation using the following theorem.

**Lemma 5.3.** *For a d-dimensional normal random variate $X \sim N_d(\mu, \Sigma)$ and a linear mapping $A : \mathbb{R}^d \to \mathbb{R}^q$ we obtain $AX \sim N(A\mu, A\Sigma A')$.*

*Proof.* See Anderson (1958, Theorem 2.4.1). □

Assuming a parametric model for a return distribution, one has to consider that the parameters usually change over the time. How to estimate these changing parameters, is one difficult problem (recall from Section 1.6 that volatility estimates have poor finite sample properties). Another point is that appropriate distributional assumptions should be made. If the parametric model is correct, the statistician is rewarded with a lower estimation risk for extreme quantiles (Huisman, Koedijk, and Pownall 1998). See also Section 1.6 for nonparametric and parametric quantile estimation results under the assumption that the returns distribution is generalized hyperbolic.

## Generalized Hyperbolic Distributions

Let us start with a general result on densities for generalized hyperbolic distributions which corresponds to Lemma 5.3.

**Theorem 5.4.** *Let $X$ be a $d$-dimensional random variable with symmetric generalized hyperbolic distribution, i.e. with $\beta = (0, \ldots, 0)'$, and let $h \in \mathbb{R}^d$ where $h \neq (0, \ldots, 0)'$. The distribution of $h'X$ is univariate generalized hyperbolic with the following parameters $\lambda^\times = \lambda$, $\alpha^\times = \alpha|h'\Delta h|^{-1/2}$, $\beta^\times = 0$, $\delta^\times = \delta|h'\Delta h|^{1/2}$ and $\mu^\times = h'\mu$.*

*Proof.* Let $h_1 \neq 0$ without loss of generality. Apply Theorem Ic) of Blæsild (1981) with

$$A = \begin{pmatrix} h_1 & h_2 & \cdots & h_d \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{5.5}$$

Then project the $d$-dimensional GH distribution onto the first coordinate using Theorem Ia). $\square$

The latter theorem may be used to calculate risk measures for a portfolio of $d$ assets with investments given by a vector $h \in \mathbb{R}^d$. As an example, we consider a portfolio consisting of three German stocks: Daimler Benz, Deutsche Bank and Thyssen from January 1, 1988 to May 24, 1994. The estimates are given in Table 4.6. We choose $h = (1, 1, 1)'$ and show the empirical density of the returns $h'x_t$ of the portfolio in Figure 5.2. Theorem 5.4 gives the corresponding densities obtained from the $d$-dimensional estimates of symmetric hyperbolic and symmetric NIG distributions. Note, that the distribution of the sum of multivariate NIG distributions is again NIG, whereas the distributon of the sum of hyperbolic distributions is generalized hyperbolic with $\lambda = (d+1)/2$. Figure 5.2 shows also the direct estimate for the univariate GH distribution to the returns of the whole portfolio. We obtain the estimate $\lambda = -1.96$ for the one-dimensional generalized hyperbolic distribution, which corresponds to similar estimates of $\lambda$ for univariate asset return distributions.

The densities and log-densities indicate that symmetric GH distributions lead to a rather precise modelling of the return distribution of the portfolio. As a consequence, one can get more realistic risk measures than the traditional ones based on the normal distribution. Figure 5.3 shows the Value-at-Risk over a 1-day horizon with respect to a level of probability $\alpha \in (0, 1)$ and the shortfall.

## Reduction to One Dimension

This approach corresponds to the approach in the middle column of Table 5.1: We reduce the multivariate portfolio to one dimension and apply univariate GH distributions. This approach is a shortcut, which allows to avoid the estimation of multivariate distributions. Moreover, in the case of a crash, the "correlation" of assets is close to unity. Therefore, multivariate distributions based on a correlation matrix can be a good model for the "center" of the distribution, but there might be problems with the modelling of extreme tails. In particular, tails of univariate distributions derived from multivariate GH laws by Theorem 5.4 can be deficient, e.g. for levels of probability below 1% (see Section 4.1, The Epps Effect).
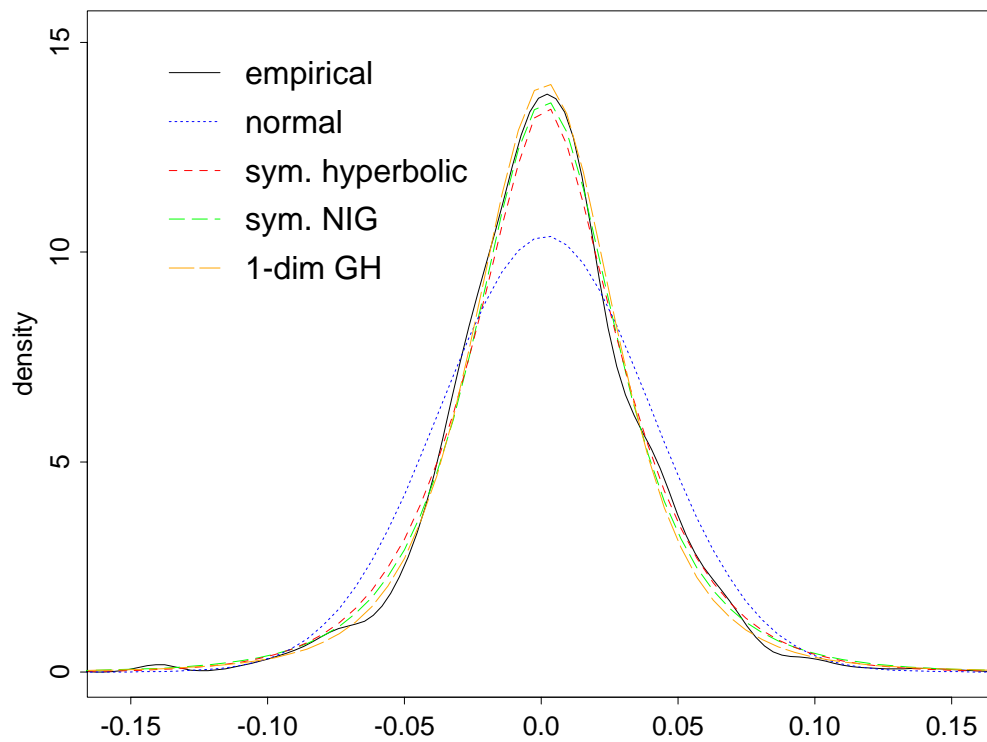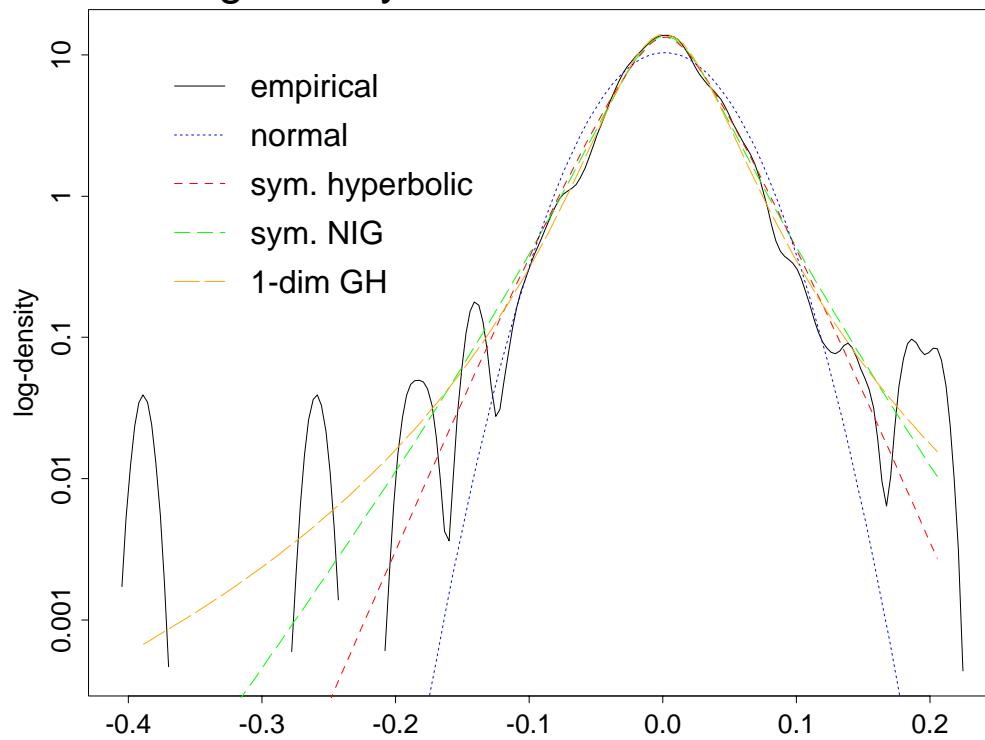
Figure 5.2: *Distribution of the returns of a portfolio consisting of Daimler Benz/Deutsche Bank/Thyssen. Investment of one DM in each asset.*

## VaR: 1 DAI + 1 DBK + 1 THY



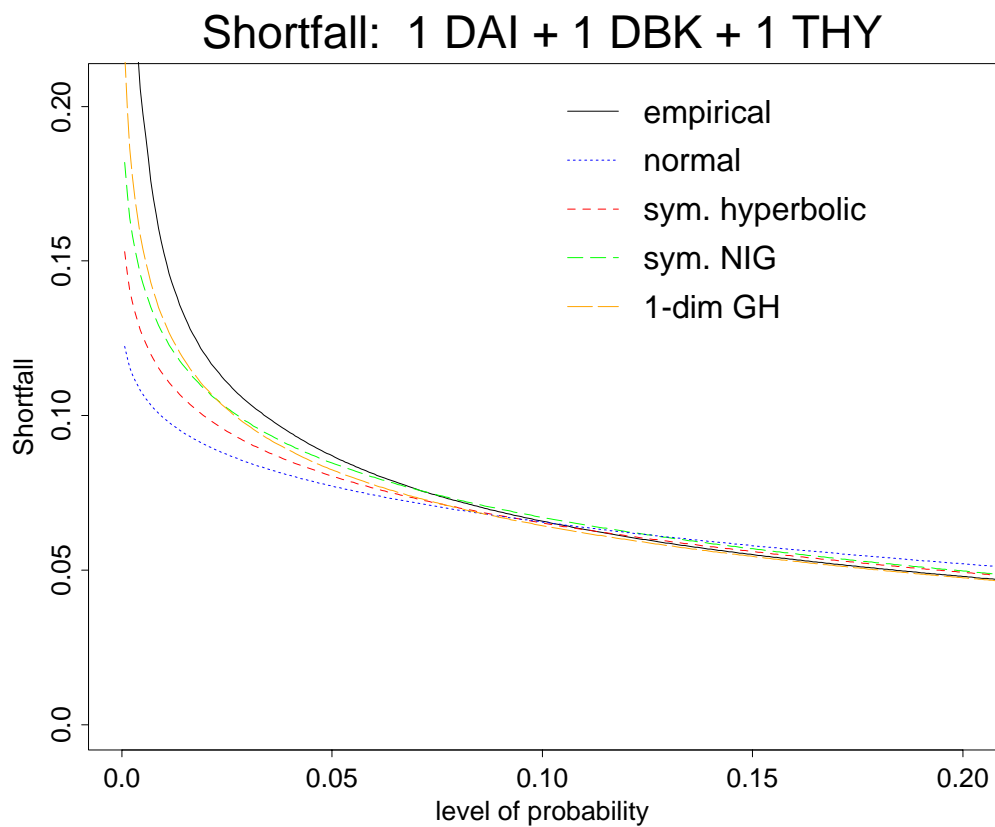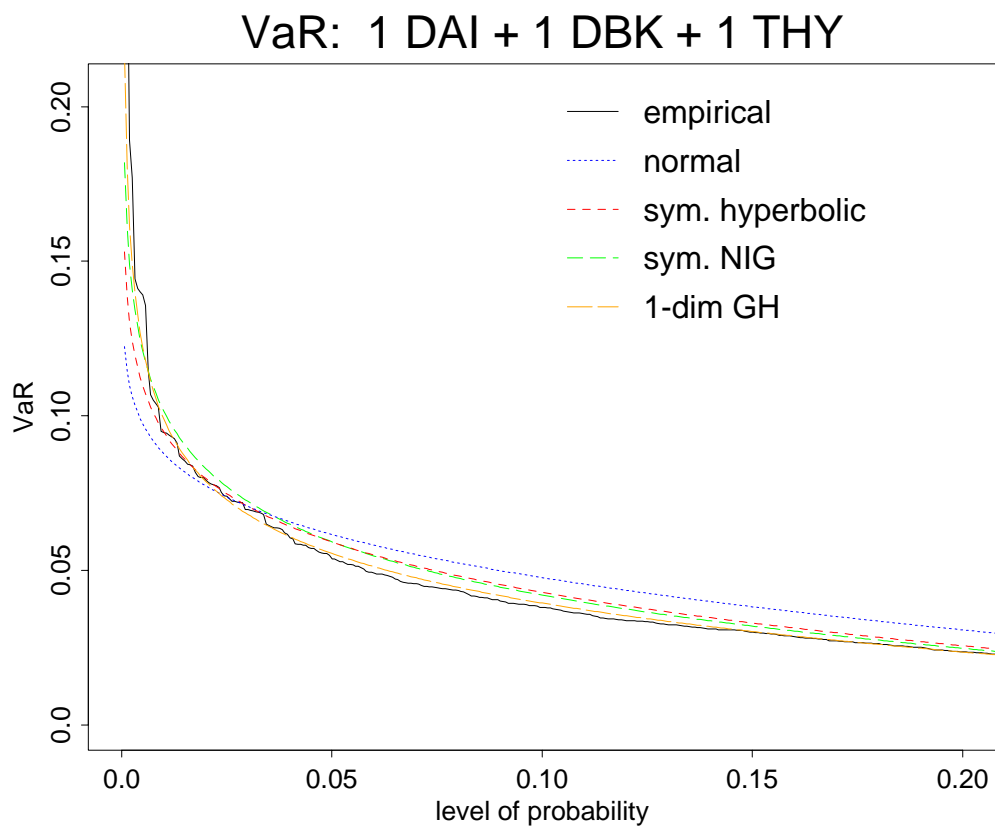## Shortfall: 1 DAI + 1 DBK + 1 THY



Figure 5.3: *Risk measures for Daimler Benz/Deutsche Bank/Thyssen.*

## Long-Term Shape Parameters

An approach similar to the rescaling mechanism proposed in the univariate case is the estimation of the shape (characterized by $(\lambda, \zeta)$) from a longer time period and to use an up-to-date covariance matrix $S$. This allows to incorporate risk of extreme events which do not occur in the preceding 250 trading days, the minimum time period proposed by the Basel Committee on Banking Supervision (1995). Therefore, we have to choose a subclass, i.e. a parameter $\lambda \in \mathbb{R}$, and to fix a long-term estimate for $\zeta$. We compute the matrix $\Sigma$ in the second parametrization by

$$\Sigma = \delta^2 \Delta = \frac{\zeta}{R_\lambda(\zeta)} S \qquad \text{and} \qquad |\Delta| = 1. \tag{5.6}$$

For numerical reasons we compute $\delta$ directly

$$\delta = \sqrt{\frac{\zeta}{R_\lambda(\zeta)}} |S|^{1/(2d)}. \tag{5.7}$$

The obtained symmetric GH distributions are again used to calculate VaR. The use of a long-term shape parameter incorporates a high possibility of extreme events, even if no crash in the preceding 250 trading days occured.

## Related Approaches: VaR-x

We also want to mention a different but closely related approach to risk measurement which is proposed by Huisman, Koedijk, and Pownall (1998): They replace the normal distribution by a (univariate) Student-t distribution and call their approach to risk measurement *VaR-x*. Indeed, it is a special case of the GH risk measurement approach examined in this study. Huisman, Koedijk, and Pownall (1998) propose to estimate Student-t distributions from one tail of the distribution using a modified Hill-estimator. The obtained tail index $\alpha$ together with the empirical mean and variance allows to determine a particular Student-t distribution. The tail index also identifies the number of moments which do exist. They have tested this method with good results.[3] However, a multivariate extension of this particular approach is not straightforward, because the tail index is determined from only one side of a univariate empirical distribution.

## 5.3 GH-IGARCH

In contrast to some standard financial models where the volatility is assumed to be constant, in the financial markets frequent changes in the "temperature" of the market are observed. In contrast, e.g. to the historical simulation, ARCH-type models allow for the necessary volatility updating. This leads to models with heavier tails than the normal distribution. We have applied the multivariate IGARCH model of Nelson (1990). In this model the variance

---

[3]They have excluded the crash of '87 from their data set. This is not appropriate in the evaluation of a risk measurement method.

$\sigma_{1,t}^2$ and covariance $\sigma_{12,t}^2$ are given by

$$\sigma_{1,t}^2 = (1 - \lambda) \sum_{t \geq 1} \lambda^{t-1}(r_t - \bar{r}) \tag{5.8}$$

$$\sigma_{12,t}^2 = (1 - \lambda) \sum_{t \geq 1} \lambda^{t-1}(r_{1,t} - \bar{r}_1)(r_{2,t} - \bar{r}_2), \tag{5.9}$$

where $\lambda$ is a decay factor, $r_t, r_{1,t}, r_{2,t}$ returns of financial assets and $\bar{r}, \bar{r}_1, \bar{r}_2$ the corresponding mean values. This model for the volatility corresponds to the computation of variance and covariance with exponential weights. Therefore, it is sometimes called *exponentially weighted moving average* (EWMA) model.

The computation of IGARCH estimates is very simple since the following recursion formulas may be applied

$$\sigma_{1,t}^2 = \lambda\sigma_{1,t-1}^2 + (1 - \lambda)r_{1,t}^2, \tag{5.10}$$

$$\sigma_{12,t}^2 = \lambda\sigma_{12,t-1}^2 + (1 - \lambda)r_{1,t}r_{2,t}. \tag{5.11}$$

Although GARCH models have heavier tails, Bollerslev (1987) mentioned that the residuals are still not normal. Another motivation for the application of the IGARCH model in this study is, that it is implemented in RiskMetrics and used as a standard tool by risk managers. To guarantee comparability, we have used the decay factor $\lambda = 0.94$, which is also applied in J.P. Morgan and Reuters (1996) for daily returns.

Now we can also take this IGARCH estimate of the variance matrix and construct a GH distribution following the procedure described in (5.6) and (5.7). We used the same long-term estimates for $\zeta$ as above. Of course, further refinements are possible (and necessary): For instance the shape parameter $\zeta$ may be estimated from residuals of Gaussian IGARCH models. However, often it is possible to estimate the parameters of the GARCH process without fixing a particular distribution of the innovations, whenever the distribution is normalized and centered. Then the distribution itself may be estimated using the residuals. To distinguish this approach from Gaussian GARCH approaches we have called it GH-IGARCH.

## 5.4 Backtesting Experiments

In the sequel we are going to compare the model-generated risk measures with actually observed market risk exposures for certain example portfolios. The quality of quantile-based risk measures like VaR is usually evaluated by backtesting, i.e. counting excess losses over an estimated quantile value. The Basel Committee on Banking Supervision (1996) stresses that "backtesting offers the best opportunity for incorporating suitable incentives into the internal models approach." However, no methodology for the backtesting has been singled out as industrial standard. See Davé and Stahl (1997) for a study focussing on the variance-covariance and GARCH approaches to estimate VaR.

We examine the methodologies to estimate VaR following the backtesting proposals of the Basel Committee on Banking Supervision (1995, Paragraph IV.23). The Basel Committee on Banking Supervision (1995, Paragraph IV.3) has chosen a holding period of 10 days for the calculation of capital requirements. First of all with the objective to put more weight on options. Nevertheless, we use a one-day horizon because our example portfolios do not include options and the increased number of nonoverlapping returns in the observation period allows

more accurate statistical results. Moreover, for backtesting in contrast to the calculation of capital requirements, the Basel Committee on Banking Supervision (1996) prefers a 1-day trading period. The Basel Committee on Banking Supervision (1996) compares VaR estimates with trading outcomes for each trading day and with changes in the price without changes in the portfolio composition. Working with a 1-day trading horizon minimizes the contamination of the results by trading. Moreover, trends and fees do have a smaller influence in shorter time intervals. In the case of this backtesting experiment these arguments are not applicable, since we do not model trading during the holding period at all. After computing VaR for each day in a time period from the preceding 250 days, we count the losses occuring on the day after the estimation period which are greater than the Value-at-Risk.[4] Since Value-at-Risk is essentially a quantile estimate, the percentage of excess losses should correspond to the level of probability $\alpha$.

We have performed the backtesting experiment for the following portfolios:

**German stocks:** 1 DAI + 1 DBK + 1 THY

**NYSE-Indices:** 1 Industrial + 1 Transport + 1 Utility + 1 Finance

**Global Indices in Deutsche Mark:** 1 dowdem + 1 ftsdem + 1 daxdem + 1 nikdem

**FX:** 1 usddem + 1 jpydem + 1 gbpdem + 1 chfdem

The notation means a fixed investment of one currency unit in each asset (not the number of assets). In the sequel we describe only the results for the first data set. The results for the other data sets are comparable to those described for the example portfolio with three German stocks.

The results given in Table 5.4 show that both standard estimators for Value-at-Risk underestimate the risk of extreme losses. This effect is visible in the percentage of excess losses in the historical simulation. In the variance-covariance approach we observe too high values for the level of probability $\alpha = 1\%$ and too small values for $\alpha = 5\%$. This corresponds to Figure 5.3 (top). The percentages of realized losses greater than VaR are closer to the level of probability in both cases, $\alpha = 1\%$ and $\alpha = 5\%$ for the symmetric hyperbolic and the symmetric NIG distribution than in the standard approaches. Even better results, especially at the 1% level, are obtained in the GH approaches with a long-term shape parameter $\zeta$ and in the GH-IGARCH methodology.

Note, that the Basel Committee on Banking Supervision (1995) preferes 99% quantiles to lower quantiles.

## Three Zones and the "Factor Three"

The market risk capital that a US bank must hold is calculated as the maximum of the current VaR estimate and the average of recent VaR estimates

$$\max\left\{\text{VaR}_t(10 \text{ days}, 1\%);\ F_t \times \frac{1}{60} \sum_{i=0}^{59} \text{VaR}_{t-i}(10 \text{ days}, 1\%)\right\} + CR_t,$$

where $F_t \geq 3$ is a multiplicative factor and $CR_t$ is an additional capital charge for the portfolios idiosyncratic credit risk (Lopez 1999). To determine $F_t$ the exceptions, i.e. the

---

[4]The Basel Committee on Banking Supervision (1996) denotes them as "exceptions".

Table 5.4: *Ex post evaluation of risk measures: percentage of losses greater than VaR. Each trading day the value at risk for a holding period of one day is estimated from the preceding 250 trading days (German stocks: Daimler Benz/Deutsche Bank/Thyssen from January 1, 1989 to May 24, 1994, Investment of one Deutsche Mark in each asset).*

| VaR methodology | % | $\alpha = 1\%$ Max Loss/VaR | $\alpha = 5\%$ % |
|---|---|---|---|
| Historical Simulation | 2.08 | 5.9295 | 5.79 |
| Variance-Covariance | 1.63 | 6.0266 | 4.45 |
| RiskMetrics / IGARCH | 1.34 | 5.8509 | 4.75 |
| Sym. hyperbolic | 1.34 | 5.3073 | 4.6 |
| Sym. NIG | 1.26 | 5.4286 | 4.75 |
| Sym. hyperbolic, long-term $\zeta$ | 1.26 | 5.3744 | 4.45 |
| Sym. NIG, long-term $\zeta$ | 1.04 | 5.1595 | 4.9 |
| Hyperbolic IGARCH, long-term $\zeta$ | 1.11 | 5.247 | 4.82 |
| NIG IGARCH, long-term $\zeta$ | 1.11 | 5.0464 | 5.34 |
| 1-dim HYP | 1.41 | 5.9214 | 4.9 |
| 1-dim NIG | 1.41 | 5.8397 | 4.97 |

returns below the (negative) VaR(1 day, 1%) in the recent 250 trading days, are counted (see also the Basel Committee on Banking Supervision 1996). Note, that for the determination of $F_t$ the VaR is calculated for a holding period of one trading day which corresponds to the results in Table 5.4. In the "green zone", i.e. with four or fewer exceptions, $F_t$ remains at the minimum value of three. From five to nine exceptions (in the "yellow zone") the factor $F_t$ increases incrementally. With ten or more exceptions in the red zone the regulatory authority requires an improvement of the risk manangement system. This underlines the necessity to have good estimates of VaR at the 1% level with a holding period of one trading day. See the Basel Committee on Banking Supervision (1996) for the statistical reasoning behind the definition of the three zones. Note, that the "green zone" ends at 1.6%, therefore, the results for the historical simulation and the variance-covariance approach in Table 5.4 would trigger a supervisor's inspection, if this models would have been used in a financial institution.

From a statistical point of view it is not clear, how to justify this factor $F_t$. In Table 5.4 we have given the maximum of the quotient of observed loss and VaR for the different risk-measurement approaches. Obviously this factor is smaller in all GH based approaches than in the three standard methods.

In Table 5.5 we give the same quotients for the other example data sets. We have chosen the variance-covariance approach to faciliate a comparison. Note, that this factor does not correspond exactly to $F_t$ as given above, but it hints at the tail distribution of the example portfolios.

Obviously these portfolios differ substantially in their return distribution below a 1% level of probability. The Value-at-Risk multiplied by a standard factor seems not to be very reliable as a basis for the calculation of the capital, which is necessary to cover crash events.

Table 5.5: *Maximum of the quotient of observed loss and VaR. The VaR estimate is calculated for a 1 trading day horizon and a 1% level of probability following the variance-covariance approach.*

| Dataset | Max Loss/VaR |
|---|---|
| German stocks: Daimler Benz, Deutsche Bank, Thyssen | 6.0266 |
| NYSE Indices: Industrial, Transport, Utility, Finance | 2.6841 |
| Indices in DEM: dowdem, ftsdem, daxdem, nikdem | 2.7957 |
| FX rates: usddem, jpydem, gbpdem, chfdem | 1.9222 |
| 20 financial assets, NMZF dataset | 2.6994 |
| 30 German stocks, DAX | 8.6444 |

## 5.5  Statistical Tests

**Evaluation of Distribution Forecasts**

In the last paragraph we have only discussed the results of the backtesting experiment on the basis of two lower quantile values, 1% and 5%. Let us also look at the whole distribution (Breckling, Eberlein, and Kokic 1999).

Essentially risk-measurement is based on distribution forecasts for a period ahead. In the backtesting experiment we obtain a prediction for the distribution of the return on a particular day. The estimated distribution is e.g. based on the preceding 250 trading days. We plug the actually observed return in the probability function of the estimated distribution and repeat this procedure for the whole data set. Then we obtain, under the assumption of independence, values which should obey a uniform distribution on $[0, 1]$. In principle, this is the inverse procedure to the generation of GH random variates: They are obtained by plugging uniform random variates into the quantile function. Figure 5.6 shows histograms of these observations. Since these observations should be uniformly distributed, values smaller than one indicate that returns in the corresponding range of quantiles, e.g. from 0% to 10%, occur less frequently than predicted by the forecasted distribution.

Basically, the plots in Figure 5.6 allow for an analysis of the overall fit of a particular approach used for risk-measurement. For instance, the normal distribution underestimates the probability of small returns, i.e. the values in the bins from 40% to 60% are too high. A more flat pattern is observed for the GH-based than for normal approaches. Nevertheless, it is obvious that the best overall fit can be observed for the historical simulation and for the reduction to a 1-dimensional dataset.

However, the underestimation of the frequency of returns in the tails is not clear form Figure 5.6. Choosing smaller classes as in Figure 5.7 shows that there are more observations in the tails than predicted by normal distributions. The values of extreme returns, in the way they are given in Table 5.4, seem to be more useful for the examination of lower quantiles.
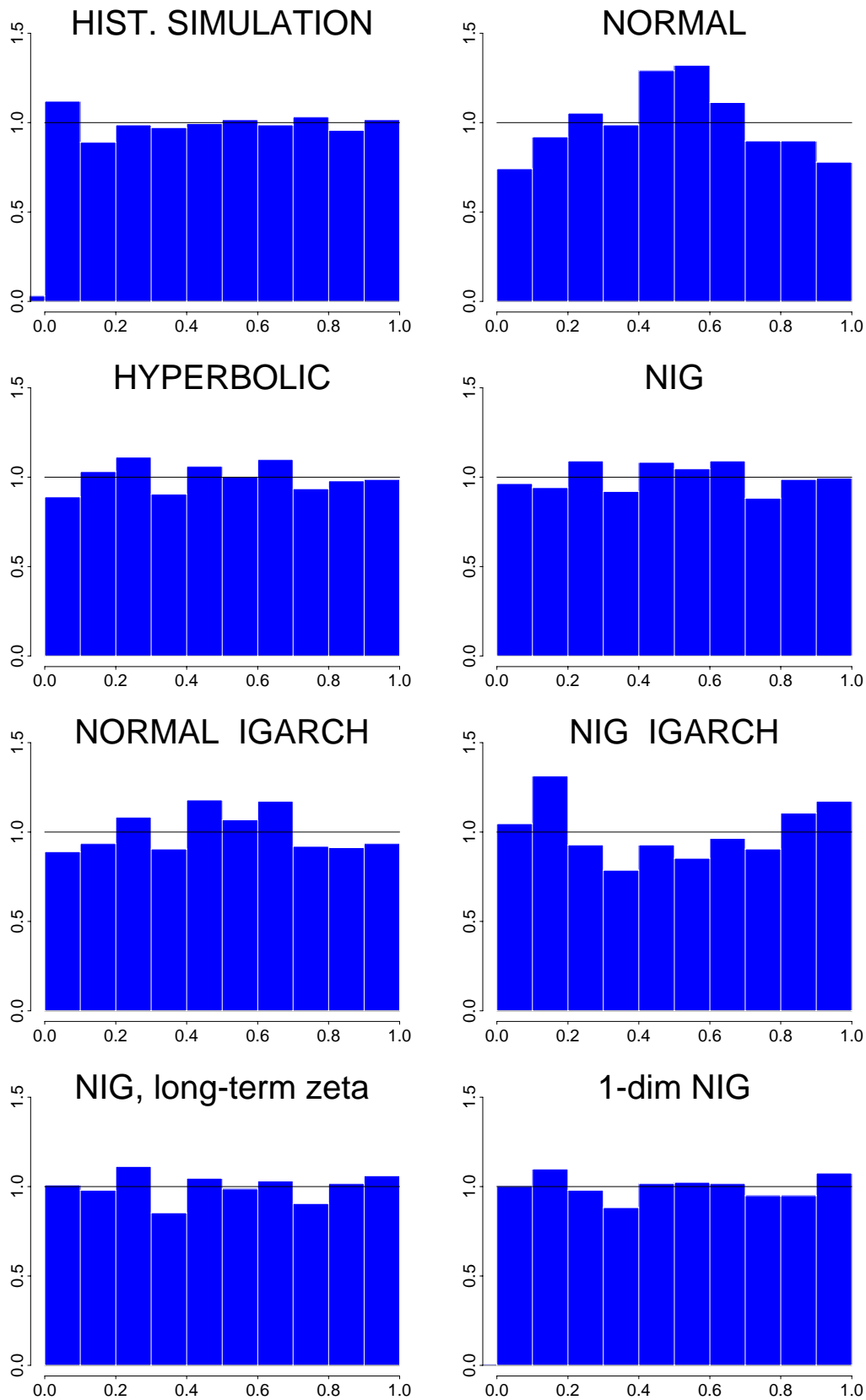
Figure 5.6: *Visualization via CDF: 1 DAI + 1 DBK + 1 THY / German stock, January 1, 1988 - May 24, 1994*
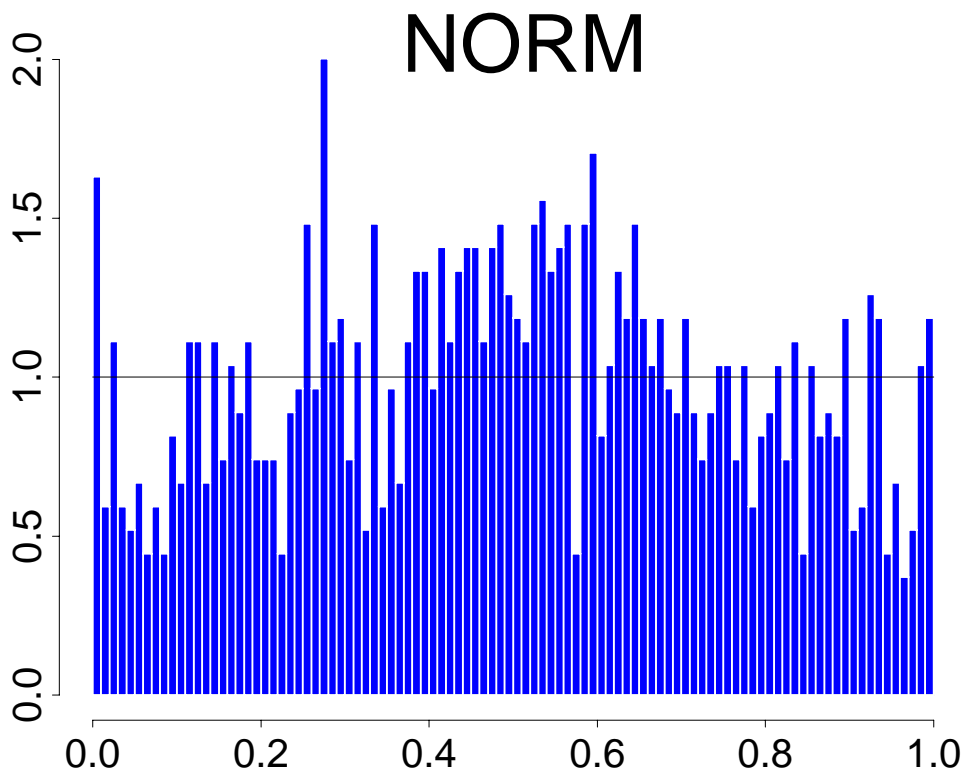
Figure 5.7: Visualization via CDF: 1 DAI + 1 DBK + 1 THY / German stock, January 1, 1988 - May 24, 1994

## Statistical Tests

The fact that the transformed observations should be uniformly distributed on $[0, 1]$ under the assumption that the model behind the risk measurement approach is correct, allows for a statistical test. We performed a Kolmogorov-Smirnov and a $\chi^2$-test under the hypothesis that the true distributions are mutually independent and uniform. The results are given in Table 5.8.

Both tests reject the hypothesis in the case of the variance-covariance approach; we also get low p-values for all IGARCH models. We obtain high p-values for the historical simulation and the reduction to a 1-dimensional data set. The hypothesis of a uniform distribution is not rejected for symmetric GH with or without long-term shape parameter $\zeta$. The ranking of hyperbolic and NIG distributions varies in these tests. Similar results are obtained for the other data sets.

## Tests of Independence of Residuals

The statistical tests in this chapter are focused on the tail behaviour and the goodness of fit of the quantile estimates. Another question is whether clustering effects in the exceedences occur or not. In particular, this allows for a comparison of the different volatility estimators. Since our research topic is the comparison of the "driving" distributions, we have not pursued this question further; see Davé and Stahl (1997, 4.3) for relevant statistical tests.

Table 5.8: *German stock, January 1, 1988 – May 24, 1994: 1 DAI + 1 DBK + 1 THY / Holding period: 1 trading day.*

|  | Kolmogorov-Smirnov Test | | $\chi^2$-Test | |
|---|---|---|---|---|
|  | KS-Statistic | p-value | $\chi^2$-Statistic | p-value |
| Historical Simulation | 0.015 | 0.912 | 12.736 | 1 |
| Variance-Covariance | 0.048 | 0.004 | 82.119 | $1.2e{-}05$ |
| Riskmetrics / IGARCH | 0.028 | 0.243 | 50.178 | 0.046 |
| Sym. Hyperbolic | 0.016 | 0.9 | 42.54 | 0.178 |
| Sym. NIG | 0.018 | 0.77 | 42.7 | 0.174 |
| Sym. Hyperbolic, long-term $\zeta$ | 0.018 | 0.803 | 44.516 | 0.13 |
| Sym. NIG, long-term $\zeta$ | 0.018 | 0.789 | 36.131 | 0.416 |
| Hyperbolic IGARCH | 0.03 | 0.191 | 53.062 | 0.026 |
| NIG IGARCH | 0.042 | 0.018 | 75.442 | $8.6e{-}05$ |
| 1-dim Hyperbolic | 0.017 | 0.849 | 31.27 | 0.649 |
| 1-dim NIG | 0.015 | 0.922 | 26.57 | 0.846 |

## 5.6   High-Dimensional Data

In the previous sections we have described estimation procedures and risk measures using a 3-dimensional example. In this section we will look at a 20-dimensional data set which is much less homogenous. The structure of the 20-dimensional NMZF data set has already been examined in Sections 4.1 and 4.2. Although the GH approach is an improvement in comparison to the normal distribution, the higher number of dimensions leads to a poor fit of likelihood estimates in the tails. Therefore, we propose two more sophisticated methods to by-pass this problem. Finally, we compare the results by a backtesting experiment based on the following example data sets:

**20 financial assets (NMZF):** usddem, gbpdem, caddem, nlgdem, chfdem, befdem, frfdem, dkkdem, nokdem, sekdem, itldem, atsdem, espdem, ptedem, jpydem, umlauf[5], daxdem, nikdem, ftsdem, dowdem;

**30 German stocks (DAX):** ALV, BAS, BAY, BHW, BMW, BVM, CBK, CON, DAI, DGS, DBC, DBK, DRB, MET, HEN, HFA, KAR, KFH, LIN, LHA, MAN, MMW, PRS, RWE, SCH, SIE, THY, VEB, VIA, VOW.

In the sequel we model an investment of one Deutsche Mark in each asset. See Appendix A.1 for a detailed description of the data sets.

---

[5] "Umlaufrendite", i.e. German interest rate.

## Tail-Emphasized Estimation

In Section 4.4 we have proposed to minimize the distance

$$\sum_{\alpha \in A} \left( q_{\text{emp}}(\alpha) - q_{\text{gh}}(\alpha, \zeta, \Sigma) \right)^2, \qquad A = \{0.01, 0.05, 0.95, 0.99\} \tag{5.12}$$

with respect to $\zeta$. The reason for the use of this (rather unusual) distance is the application to measure risk. Risk managers need good estimates for 1% and 5% quantiles. In contrast to other distances, e.g. the Kulback-Leibler or the Kolmogorov distance, this estimation method enforces good results for some quantiles. From a numerical point of view the implementation is rather simple, since we only have to perform a one-dimensional optimization. In Figure 5.9 we have shown the obtained errors in the calculation of VaR, i.e. errors for all levels of probability.
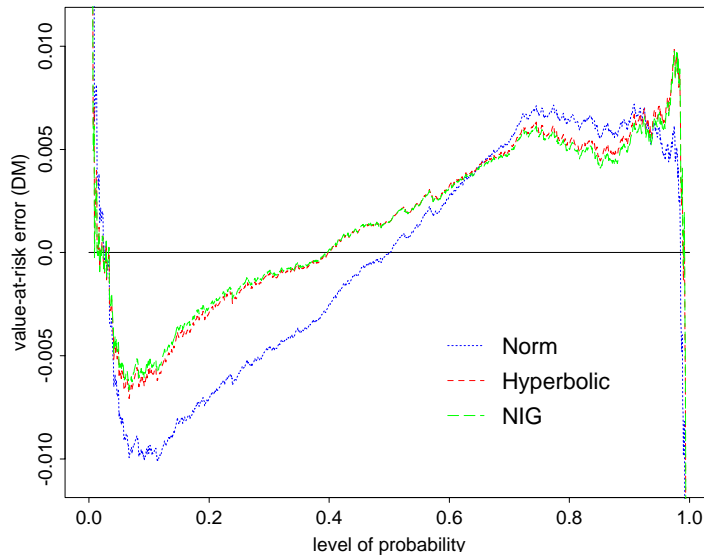


Figure 5.9: *VaR errors for normal and generalized hyperbolic distributions. The GH distributions were estimated "tail-emphasized" by minimizing (5.12).*

## Dimension Reduction

Principal component analysis allows a deeper understanding of the risk inherent in a financial market: The complexity of the market is reduced to some factors and using principal component analysis (PCA) shows the primary sources of risk to which a portfolio is exposed. See Section 4.1 for a detailed description of the PCA.

We have rescaled the vector $h_{\text{red}}$, such that the variance of the whole portfolio is constant

$$\text{Var}[h'X] = \text{Var}[h'_{\text{red}}Y],$$

where $X$ is the original and $Y$ the reduced random vector. The results of this approach applied to a 20-dimensional data set are shown in Figure 5.10. The empirical and the "reduced" empirical distribution and VaR estimates respectively are close. Therefore, not much

explanatory power is lost by the reduction to 3 dimensions. Obviously GH distributions are closer to the empirical and the "reduced" empirical distribution in the center and also in the tails. In the backtesting experiment we will also consider the combination of this dimension reduction method with the tail-emphasized estimation of GH parameters.

See Christiansen (1999) for a similar approach to apply dimension reduction to the modelling of ARCH time series. Note, that financial institutions use PCA to construct stress testing scenarios (Loretan 1997, 3.3). For the backtesting experiment we have chosen a reduction to three factors, which is of course somewhat arbitrary. In contrast to the other approaches to risk measurement we observe a striking difference between the 20-dimensional mixed data set of financial assets and the 30 German stocks for the VaR approaches based on PCA.

## Backtesting

To summarize the results of the high-dimensional backtesting experiment, we emphasize the following as results from Table 5.11:

- Comparing the results for both data sets, the 20 mixed assets and the 30 German stocks, it is clear, that VaR estimation works better for the German stocks. This data set is more homogeneous, see for instance Section 4.1.

- For the 20 dimensional data set we always observe an underestimation of risk at both levels of probability. With respect to the 1% VaR in the case of the 30 German stocks, the risk is only overestimated in one NIG approach. At the 5% level we observe both, over- and underestimation, for the 30 German stocks. Looking at the VaR plots, e.g. in Figure 5.3, one expects these mixed results for the normal distribution.

- Historical Simulation: We observe a substantial underestimation of risk. With exception of the 1% VaR for the 20 mixed assets, the historical simulation yields the worst results of all analyzed approaches. In view of the poor finite-sample properties of nonparametric quantile estimators (see Section 1.6) this is not surprising. A longer estimation period may improve the ability of the historical simulation to cope with large returns. However, this leads to volatilities which are less up-to-date.

- Variance-Covariance: As expected (see Figure 5.3), we observe a striking underestimation of risk at the 1% level of probability and better results at the 5% level.

- RiskMetrics/IGARCH: This approach allows for an improvement of the variance-covariance approach in most cases. Nevertheless, the underestimation of risk at the 1% level in the normal model confirms the remark of Bollerslev (1987), that residuals of ARCH-type models are still not normal.

- Application of symmetric GH distributions leads to an improvement of backtesting results in comparison to the historical simulation and corresponding apporaches based on normal distributions. The only exception are symmetric hyperbolic distributions estimated from the preceeding 250 trading days.

- Long-term shape parameter: The application of a long-term estimate for the shape parameter $\zeta$ improves the VaR estimation substantially.
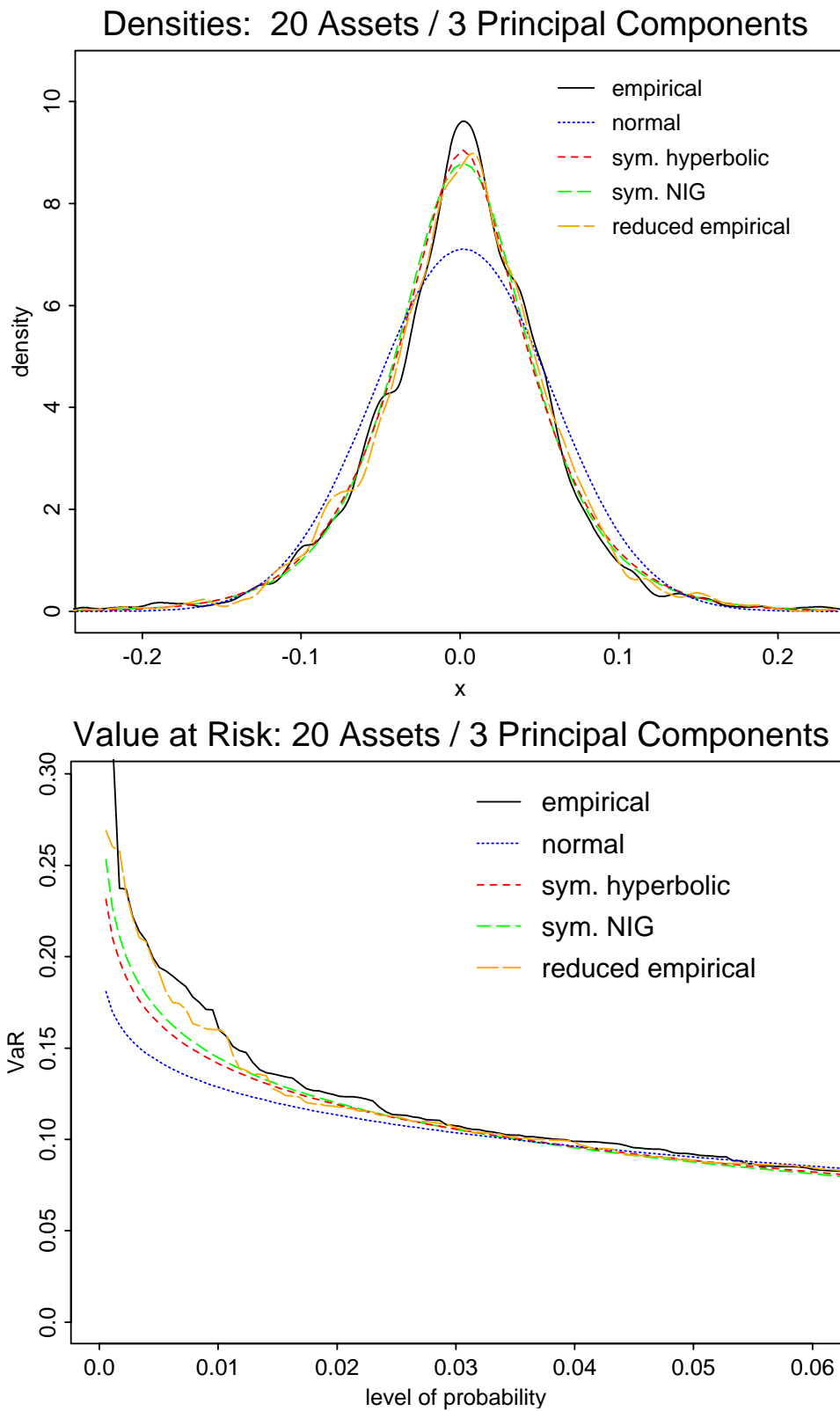
Figure 5.10: *Distribution of 20 assets (NMZF data set) from the point of view of a German investor: usddem gbpdem caddem nlgdem chfdem befdem frfdem dkkdem nokdem sekdem itldem atsdem espdem ptedem jpydem umlauf daxdem nikdem ftsdem dowdem (Investment of 1 DM in each asset).*

Table 5.11: *Ex post evaluation of VaR measures applied to high-dimensional data sets (holding period one trading day, fixed investment of one currency unit in each asset). The "green zone" ends at 1.6% for the 1% level.*

| | 20 financial assets | | 30 German stocks | |
|---|---|---|---|---|
| VaR methodology | 1% | 5% | 1% | 5% |
| Historical Simulation | 2.02 | 6.01 | 1.78 | 6.75 |
| Variance-Covariance | 2.42 | 5.55 | 1.56 | 4.08 |
| RiskMetrics / IGARCH | 2.02 | 5.23 | 1.63 | 5.27 |
| *GH-based approaches* | | | | |
| Sym. hyperbolic | 2.42 | 5.81 | 1.56 | 4.38 |
| Sym. NIG | 1.70 | 5.68 | 1.41 | 4.30 |
| Sym. hyperbolic long-term $\zeta$ | 2.22 | 5.62 | 1.48 | 4.23 |
| Sym. NIG long-term $\zeta$ | 1.44 | 5.75 | 1.26 | 4.38 |
| Hyperbolic IGARCH | 2.02 | 5.23 | 1.63 | 5.27 |
| NIG IGARCH | 1.63 | 5.42 | 1.11 | 5.42 |
| *Tail-emphasized estimation of the shape parameter* | | | | |
| Sym. hyperbolic | 2.29 | 5.62 | 1.48 | 4.15 |
| Sym. NIG | 1.89 | 5.68 | 1.26 | 4.38 |
| Sym. hyperbolic long-term $\zeta$ | 2.22 | 5.62 | 1.56 | 4.08 |
| Sym. NIG long-term $\zeta$ | 1.44 | 5.75 | 1.56 | 4.15 |
| Hyperbolic IGARCH | 2.02 | 5.23 | 1.63 | 5.27 |
| NIG IGARCH | 1.63 | 5.42 | 1.63 | 5.27 |
| *Dimension reduction using PCA, three Factors* | | | | |
| Historical Simulation - reduced | 2.02 | 6.14 | 1.71 | 7.20 |
| Variance-Covariance | 2.68 | 5.49 | 1.34 | 4.53 |
| Sym. hyperbolic | 1.63 | 5.55 | 1.11 | 5.27 |
| Sym. NIG | 2.16 | 5.88 | 0.89 | 5.34 |
| Sym. hyperbolic, tail-emph. est. | 2.09 | 5.49 | 1.11 | 4.60 |
| Sym. NIG, tail-emph. est. | 2.09 | 5.62 | 1.11 | 4.67 |
| *Reduction to one dimension* | | | | |
| Hyperbolic | 1.37 | 5.29 | 1.19 | 4.97 |
| NIG | 1.31 | 5.49 | 1.11 | 5.04 |

- GH-IGARCH: The results are superior to the normal IGARCH approach. Note, that we have applied the shape parameter $\zeta$ estimated under the assumpiton of independence—better results may be obtained for an appropriately estimated $\zeta$.

- Normal inverse Gaussian distributions are always superior at the 1% level to the corresponding approaches with hyperbolic distributions.

  This is in accordance with the estimation results in Table 4.6, where NIG distributions have a higher likelihood or a lower squared distance of quantiles respectively.

- Tail emphasized estimation: We observe an improvement to the calculation of VaR based on likelihood estimates for the less sophisticated approaches (Sym. hyperbolic and Sym. NIG) whereas the differences in the long-term $\zeta$ and in the GH-IGARCH cases are only small and unsystematic.

- Dimension reduction by PCA: The reduction of dimensions to three factors yield a further improvement to standard an GH based risk-management approaches in the case of the 30 German stocks. The results in the case of the 20 mixed assets are not that good.

  From Figure 4.2 it is clear that the same number of factors explains only a smaller portion of the variance in the case of the 20 dimensional data set in comparison to the more homogenous 30 German stock data set. Unfortunately, an increase in the number of factors for the 20 financial assets does not change the results substantially. Recall, that there is no canonical way to choose the number of factors.

  The principal advantage of this approach is, that it allows to get better insights into the origins of market risk.

- Combination of dimension reduction and tail-emphasized estimation: The results are mixed, no substantial improvement to the dimension reduction with likelihood estimation is observed.

- Reduction to one dimension: This approach corresponds to the approach in the middle column of Table 5.1. Obviously we obtain the best results of all approaches with this method. In particular, the experiment yields the best results for the 20-dimensional data set and the 1% level of probability.

Summarizing the results, using the NIG distribution for high-dimensional data sets in conjunction with a long-term shape parameter $\zeta$, which may be obtained via tail-emphasized estimation and an up-to-date covariance matrix seems to be a reasonable approach from a statistical and numerical point of view. Of course, it is favourable to estimate the covariance matrix also following the IGARCH approach. Stochastic volatility models like GARCH have the potential to avoid serial correlation in the exceedances—a topic which we have not examined further in this study. Because of the recursion formulae (5.10) and (5.11) the NIG-IGARCH approach is computationally not demanding. Since NIG distributions are closed under convolution, it is also straightforward to compute the 10-day distributional forecast from one-day forecasts.
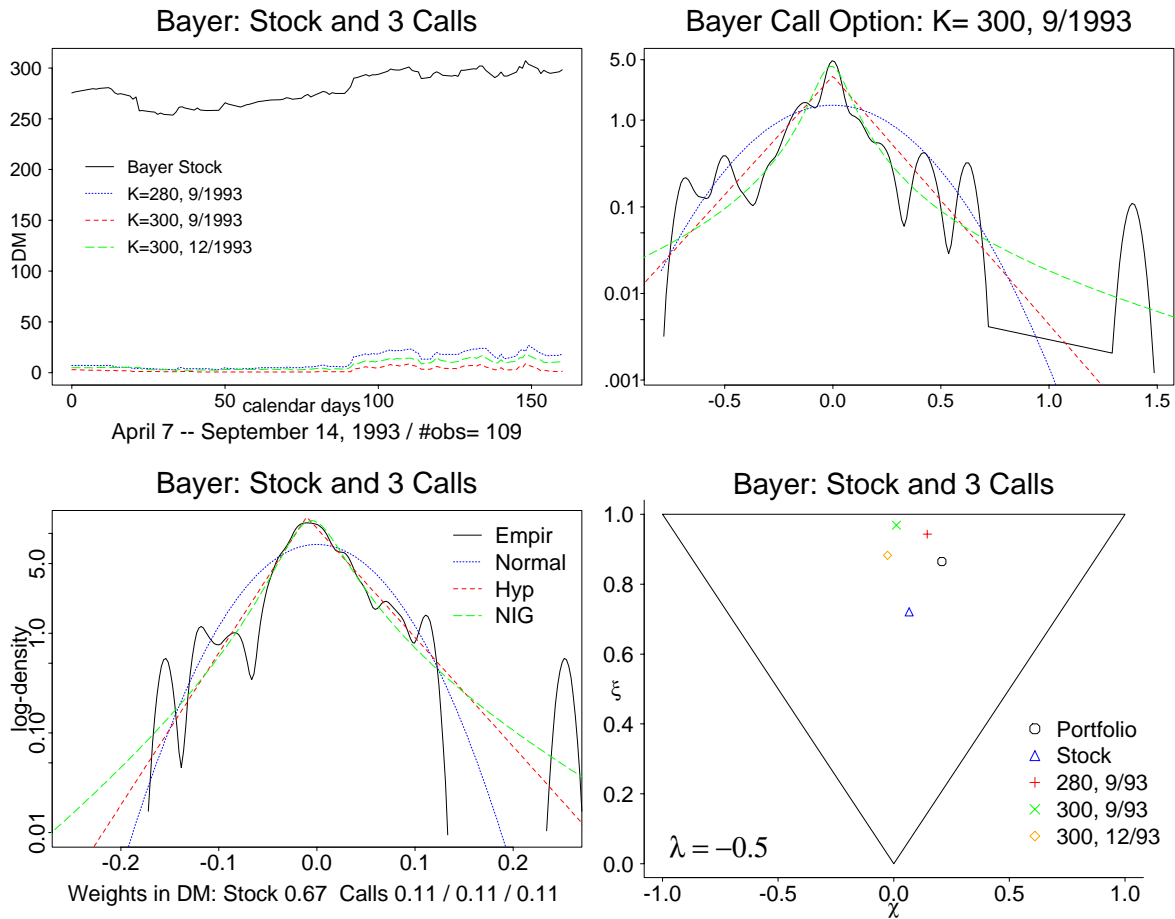
Figure 5.12: *Bayer stock and three call options, April 7 – September 14, 1993.*

## 5.7   Nonlinear Portfolios, GH Simulation Method

In view of the leverage effect of options it is necessary to pay particular attention to nonlinear portfolios and to develop appropriate statistical methods.

First insights into the behaviour of portfolios with nonlinear risk are obtained by looking directly at the return distribution of an example portfolio with options. Figure 5.12 shows the paths of the stock and the three options (top left), the distributions of returns of one call (top right) and the whole nonlinear portfolio (bottom left). The last graph contains the shapes of estimated NIG distributions. Obviously, as a result of the leverage effect, return distributions of derivatives have a higher kurtosis and they live on a different scale. Furthermore, the NIG distribution estimated for the portfolio is skewed in contrast to NIG distributions estimated for linear portfolios. See the NIG shape triangle for the 30 German stocks in the DAX in Barndorff-Nielsen and Prause (1999).

The calculation of risk measures is often based on a linear approximation. For instance, the latter approach is used for assets with currency risk or for portfolios with options (Bühler, Korn, and Schmidt 1998). One drawback of the linear approximation is that the leverage effect of derivative assets is often not modelled in a sufficiently realistic way. This may lead to a fatal underestimation of risk.

140

We propose to use a full-valuation approach to calculate risk-measures. Wirth (1998) has shown that GIG distributions provide a realistic model for the marginal distribution of the VDAX, the volatility index of the former Deutsche Börse AG, Frankfurt. Therefore, mixing GIG distributions can be used as an estimate for the volatility index of the market. In Chapter 3 we have mentioned some approaches to construct stochastic processes with an adequate correlation structure and GIG marginal distribution. This allows for a further refinement of our approach. In the iid case one can sample $w \sim$ GIG easily since the distribution is univaritate. The product $w\Delta$ is then used as a covariance matrix to generate normal random variates describing the asset returns. This mixing yields multivariate GH random variates. Consequently, we obtain a realistic return sample and the corresponding volatility sample. Large random samples can be obtained efficiently and be used to simulate the return distribution of nonlinear portfolios. This allows to compute the return distribution of a portfolio that includes derivatives in a fast and more accurate way than with a linear approximation. Note, that the right choice of the option pricing model is also important (Eberlein, Keller, and Prause 1998): The *model risk* caused by the choice of an unrealistic option pricing model is often neglected.

Another important aspect of the Monte-Carlo simulation of nonlinear portfolios is, that it allows to overcome the linear approximation. The advantages are stressed in the following example: Assume that a trader builds up a straddle portfolio by writing a call with strike $K_2 > S_0$ and writing a put with strike $K_1 < S_0$, where $S_0$ is the current asset price. This results in a terminal payoff function linear in the neighbourhood of $S_0$. See Figure 5.13 for the payoff function of this portfolio.
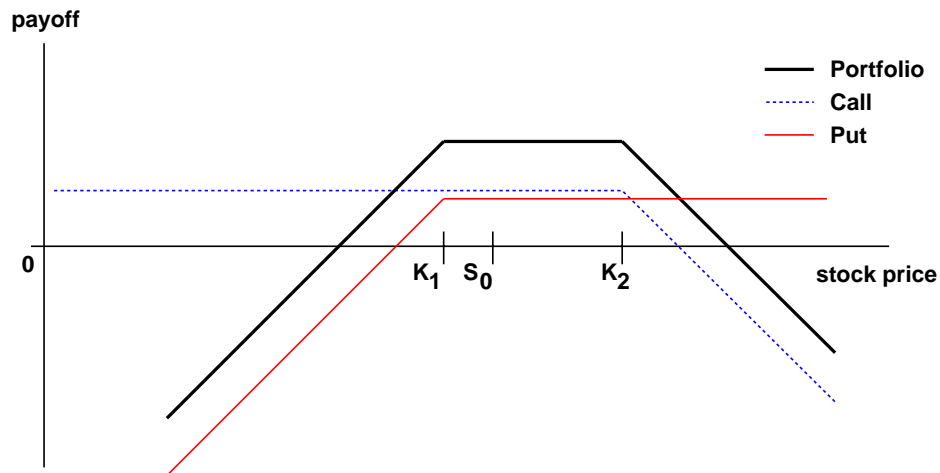


Figure 5.13: *Payoff function of a straddle.*

Large changes in the value of the asset price lead to possibly unbounded losses. These risks are not recognizable with standard methods based on linear approximations. The *GH simulation method* described above allows to forecast the return distribution of the whole portfolio in a more exact way, which is not limited by the assumption of linearity. A portfolio with a similar payoff function led to the bankruptcy of the Barings Bank.

Although unethical behaviour may always lead to severe problems, appropriate mathematical methods may prevent that a single trader risks the capital of a whole bank (see Rawnsley 1995, and Jorion 1997, Section 2.2.1 for further details). Moreover, the proposed GH simulation approach takes also the risk of frequently observed volatility changes into account. These changes are potentially hazardous if the portfolio includes derivatives and is to be sold before expiration.

## 5.8 Conclusion

In the last chapter we have applied generalized hyperbolic distributions to model risk inherent in financial markets. GH distributions improve markedly the modelling of risk. Moreover, the application is only a little more computer-intensive than models based on normal distributions.

To implement risk-measurement systems based on GH distributions one should choose a subclass and may use the algorithm proposed in Section 4.4 to estimate parameters. By Lemma 5.4 the computation of risk-measures for linear portfolios reduces to the univariate case and for nonlinear portfolios the GH simulation method based on the mixing GIG distribution (see Section 4.6) is straightforward. Sophisticated volatility estimation methods are easy to incorporate by replacing the variance estimate in (5.6), (5.7) or (4.19).

# Appendix A

# Data Sets and Computational Aspects

## A.1   Data Sets

The empirical analysis is based on the following data sets:

**DTB**  The tests of option pricing models were based on intraday data sets of options from the Deutsche Terminbörse Frankfurt (since autumn 1998 Eurex Germany). The processing of the data is described in detail in Section 2.1. See also `www.exchange.de`.

**FWB**  Returns of stock and stock indices were calculated from daily data of the Frankfurter Wertpapierbörse. We also use intraday data from the FWB to synchronize the prices of the option and the underlying. See Section 2.1 and `www.exchange.de`.

**HFDF96** Olsen & Associates (1998) released high frequency data in preparation for the Second International Conference on High Frequency Data in Finance (HFDF II), held in April 1998.

The emphasis of the second conference was however on multivariate high frequency effects and hence the data set released consists of 25 different FX spot rates, 4 spot metal rates, 12 series of Eurofutures contracts, and 2 series of indices.

These data were collected by Olsen & Associates in the period from January 1, 1996 GMT to December 31, 1996 GMT using O&A proprietary real-time data collection software. More information is available on the www-page `www.olsen.ch`.

**NMZF**  The data sets are available under the address
`www-public.tu-bs.de:8080/∼y0003876/` (`nmzf@gmx.de`).

**NYSE Indices**  The indices consist of a Composite Index of all common stocks listed on the NYSE and four subgroup indices: Industrial, Transportation, Utility, and Finance. They provide a comprehensive measure of the market trend for the benefit of many investors who are concerned with general stock market price movements.

The indices are basically a measure of the changes in aggregate market value of NYSE common stocks, adjusted to eliminate the effects of capitalization changes, new listings and delistings. The market value of each stock is obtained by multiplying its price per

143

share by the number of shares listed. The aggregate market value, which is the sum of the individual market values, is then expressed relative to a base point market value. The base value was set at 50.00 on December 31, 1965 because this figure was reasonably close to the actual average price of all common stocks at that time.

More information is available from `www.nyse.com`.

## A.2  Computational Aspects

In the following section we list the algorithms used for the numerical evaluation of the various models, densities and formulae. We explain the methodology for estimation and numerical evaluation in the text at the place, where the problem arises. We do not include all the computer codes, because it would multiply the number of pages. A German documentation of the programs is given in the manual Eberlein, Ehret, Lübke, Özkan, Prause, Raible, Wirth, and Wiesendorfer Zahn (1998) *Freiburg Financial Data Tools*.

### Univariate Model

- Densities, probability and quantile functions, moments for the GIG, GH, NIG, and hyperbolic distribution.

- Estimation procedures for the univariate generalized hyperbolic distribution:

  - for fixed and arbitrary $\lambda$. In the case of the hyperbolic and NIG distributions the simpler representations of the GH densities are used to increase the speed of the algorithms.
  - for different metrics: $L^1$, $L^2$, Kolmogorov distance and Anderson & Darling metric.

- Calculation of prices in the generalized hyperbolic and the hyperbolic model by use of Fast Fourier Transforms. Prices in the NIG case are computed using the convolution semigroup property. A description of each step in the calculation is given in Section 1.12.

- Inversion of pricing formulas with respect to volatility by the Newton iteration resp. refined bracketing methods.

- Algorithms for the estimation of implicit, historical and realized volatility in rolling windows and application of these "vola" estimates to intraday option data.

We want to remark that most algorithms are implemented efficiently to allow interactive use of the GH model. This is especially demonstrated by the hyperbolic option calculator available under `www.fdm.uni-freiburg.de/UK/`.

### Ornstein-Uhlenbeck type Volatility Model

- Estimation of the autocorrelation parameter,

- Simulation of integrated volatilities,

- Prices in the stockprice model with a volatility process of IG-OU type and superpositions of independent IG-OU type volatility processes:

- by saddlepoint approximation,

- and by integration over the simulated integrated volatility (also for Gamma-OU type processes).

- Estimation of the statistical martingale measure for the IG-OU and Gamma-OU volatility model.

## Multivariate Model

- Calculation and tests for the multivariate skewness and kurtosis introduced by Mardia (1970).

- Estimation algorithms for symmetric multivariate GH distributions:

  - Pseudo likelihood estimation for fixed $\lambda$. In the case of hyperbolic and NIG distributions the simpler versions of the density are used.

  - Tail emphasized estimation.

- Risk measures for generalized hyperbolic distributions and those for the standard approaches.

- Backtesting of risk measures.

The algorithms for the multivariate GH model are compiled in a demonstration program to show that this approach allows a quick estimation in particular for high-dimensional financial data. (See also Prause (1999b).)

## GH Package

Some of the C functions for the GH model are also collected in a package which will be part of a larger program for financial analyses (produced by Insiders GmbH, Wissensbasierte Systeme, Wilhelm-Theodor-Römheld-Straße 32, D–55130 Mainz).

## Data Handling and Visualization

A careful design of classes in Splus leads to a significant reduction of complexity in the programs. Especially the various parametrizations of the generalized hyperbolic distribution can be handled in an intuitive way.

In the context of an object-oriented language "methods" are general functions which can be applied to a particular object without bothering about details of the implementation. Consider a user who wants to plot the GH density estimated for a particular return data set. The estimates of the distribution are saved in an Splus object, e.g. `returns.gh` which is an object of the `gh`-class. The command to plot the density is simply `plot(returns.gh)`. The computer chooses the appropriate algorithm for the method `plot` and and an object of the the class `gh`. The algorithm itself usually contrasts to the simplicity of the command. In the example above, the programm has to call C functions to calculate the density and to apply adapted graphical routines. This concept is also called "information hiding", i.e. a user need not to know all tedious technical details.

- S-class `gh` for the multivariate generalized hyperbolic distribution. The object-oriented approach in Splus allows to construct a class which contains univariate as well as multivariate generalized hyperbolic distributions. The NIG, GH, hyperboloid and hyperbolic distributions are contained as special cases in an unified framework.

- Splus data structures based on the object oriented approach in Splus for intraday stock and intraday option data sets and for daily time series.

- Visualisation tools for arbitrary parameters in option data sets. We rely on the same object-oriented approach described above.

Last but not least we want to remark, that all data structures and algorithms are Y2K compliant.

# Appendix B

# Modified Bessel Functions

We summarize some properties of modified Bessel functions which are useful for the work with generalized hyperbolic distributions.

**Definition B.1.** $K_\lambda(x)$ and $I_{\pm\lambda}(x)$, $x \in \mathbb{C}$ are solutions of the differential equation

$$x^2 \frac{d^2 w}{dx^2} + x \frac{dw}{dx} - (x^2 + \lambda^2)w = 0, \tag{B.2}$$

**Remark B.3.** $K_\lambda(x)$ and $I_{\pm\lambda}(x)$ are regular function of $x \in \mathbb{C}$ throughout the x-plane cut along the negative real axis and for fixed $x \neq 0$ an entire function. $K_\lambda(x)$ tends to zero as $|x| \to \infty$ in the sector $|\arg x| < \pi/2$ and for all $\lambda$. $K_\lambda(x)$ and $I_\lambda(x)$ are real and positive when $\lambda > -1$ and $x > 0$.

*Reference.* Abramowitz and Stegun (1968, Equation 9.6.1) □

**Theorem B.4 (Integral Pepresentation).**

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty y^{\lambda-1} \exp\left(-\frac{x}{2}(y + y^{-1})\right) dy, \quad x > 0 \tag{B.5}$$

*Reference.* Barndorff-Nielsen and Blæsild (1981, App. 1) □

**Theorem B.6 (Basic Properties).**

$$K_\lambda(x) = K_{-\lambda}(x) \tag{B.7}$$

$$K_{\lambda+1}(x) = \frac{2\lambda}{x} K_\lambda(x) + K_{\lambda-1}(x) \tag{B.8}$$

$$K_{\lambda+\varepsilon} > K_\lambda(x) \qquad for\ \lambda > 0, \varepsilon > 0, x > 0 \tag{B.9}$$

*Reference.* Barndorff-Nielsen and Blæsild (1981, Equations A1.1-A1.3) and Lorch (1967, Inequality B) □

**Theorem B.10 (Relation of $K_\lambda$ and $I_\lambda$, Asymptotic Properties).**

$$K_\lambda(x) = \frac{\pi}{2} \frac{1}{\sin(\pi\lambda)} (I_{-\lambda}(x) - I_\lambda(x)) \tag{B.11}$$

$$K_\lambda(x) \sim \Gamma(\lambda)2^{\lambda-1}x^{-\lambda} \qquad as\ x \downarrow 0 \tag{B.12}$$

$$K_0(x) \sim -\ln x \qquad as\ x \downarrow 0 \tag{B.13}$$

$$\frac{K_\lambda(x)}{K_{\lambda+\varepsilon}(x)} \downarrow 1 \qquad for\ 0 < x \uparrow \infty, \varepsilon > 0, \lambda \geq 0 \tag{B.14}$$

*Reference.* Barndorff-Nielsen and Blæsild (1981) and Lorch (1967, Equation 14) □

**Theorem B.15 (Series Representation for $\lambda = n + {}^1\!/_2$, $n \in \mathbb{N}$).**

$$\mathrm{K}_{n+{}^1\!/_2}(x) \;\; = \;\; \sqrt{\frac{\pi}{2}}\, x^{-1/2}\, e^{-x} \left( 1 + \sum_{i=1}^{n} \frac{(n+i)!}{(n-i)!\, i!}(2x)^{-i} \right) \tag{B.16}$$

$$\mathrm{K}_{-1/2}(x) \;\; = \;\; \mathrm{K}_{1/2}(x) = \sqrt{\frac{\pi}{2}}\, x^{-1/2}\, e^{-x} \tag{B.17}$$

$$\mathrm{K}_{3/2}(x) \;\; = \;\; \sqrt{\frac{\pi}{2}}\, x^{-{}^1\!/_2}\, e^{-x} \left( 1 + \frac{1}{x} \right) \tag{B.18}$$

*Reference.* Barndorff-Nielsen and Blæsild (1981, Equation A1.5) □

**Theorem B.19.** *For $t > 0$ and $\lambda \geq 0$ holds*

$$\frac{\mathrm{K}_{\lambda-1}(x)}{x\,\mathrm{K}_{\lambda}(x)} \;\; = \;\; \int_0^{\infty} \frac{g_{\lambda}(t)}{x^2 + t}$$

$$g_{\lambda}(t) \;\; = \;\; \frac{2}{\pi^2 t (J_{\lambda}^2(\sqrt{t}) + Y_{\lambda}^2(\sqrt{t}))} > 0$$

*where $J_{\lambda}$ and $Y_{\lambda}$ are Bessel functions of first and second kind.*

*Reference.* Grosswald (1976, Corollary 4). □

**Theorem B.20 (Derivatives).**

$$\mathrm{K}_0'(x) \;\; = \;\; -\mathrm{K}_1(x) \tag{B.21}$$

$$\mathrm{K}_{\lambda}'(x) \;\; = \;\; -\frac{1}{2} \left( \mathrm{K}_{\lambda+1}(x) + \mathrm{K}_{\lambda-1}(x) \right) \tag{B.22}$$

$$= \;\; -\frac{\lambda}{x} \mathrm{K}_{\lambda}(x) - \mathrm{K}_{\lambda-1}(x) \tag{B.23}$$

$$(\ln K_{\lambda}(x))' \;\; = \;\; \frac{\lambda}{x} - \mathrm{R}_{\lambda}(x) \tag{B.24}$$

$$(\ln \mathrm{K}_{\lambda}(x))'' \;\; = \;\; \mathrm{S}_{\lambda}(x) - \frac{\mathrm{R}_{\lambda}(x)}{x} - \frac{\lambda}{x^2} \tag{B.25}$$

*Reference.* Abramowitz and Stegun (1968, Equation 9.6.27) and Barndorff-Nielsen and Blæsild (1981, Equations A1.3, A1.16, A1.17) □

**Derivatives with respect to Order**

*Reference.* Abramowitz and Stegun (1968, Equations 9.6.42-46) □

**Definition B.26 ($\mathrm{R}_{\lambda}$ and $\mathrm{S}_{\lambda}$).**

$$\mathrm{R}_{\lambda}(x) \;\; := \;\; \frac{\mathrm{K}_{\lambda+1}(x)}{\mathrm{K}_{\lambda}(x)}, \quad x > 0 \tag{B.27}$$

$$\mathrm{S}_{\lambda}(x) \;\; := \;\; \frac{\mathrm{K}_{\lambda+2}(x)\,\mathrm{K}_{\lambda}(x) - \mathrm{K}_{\lambda+1}^2(x)}{\mathrm{K}_{\lambda}^2(x)}, \quad x > 0 \tag{B.28}$$

*Reference.* Barndorff-Nielsen and Blæsild (1981, App. 1) ☐

**Theorem B.29 (Properties of $R_\lambda$ und $S_\lambda$).**

$$R_{-\lambda}(x) = R_{\lambda-1}(x)^{-1} \tag{B.30}$$

$$R_\lambda(x) = \frac{2\lambda}{x} + R_{-\lambda}(x) \tag{B.31}$$

$$R'_\lambda(x) = \frac{R_\lambda(x)}{x} - S_\lambda(x) \tag{B.32}$$

$$R_{-1/2}(x) = 1; \qquad R_{1/2}(x) = 1 + \frac{1}{x}; \qquad R_{-3/2} = \frac{x}{x+1} \tag{B.33}$$

$$R_{-\lambda}(x) = \frac{1}{R_{\lambda-1}(x)} = \frac{2}{\pi^2} \int_0^\infty \frac{x\,dy}{y(J_\lambda^2(\sqrt{y}) + Y_\lambda^2(\sqrt{y}))(y+x^2)} \tag{B.34}$$
$$\text{for } \lambda \geq 0, x > 0$$

$$\lim_{x\downarrow 0} R_\lambda(x) = \begin{cases} \infty & \text{if } \lambda > -1/2 \\ 1 & \text{if } \lambda = -1/2 \\ 0 & \text{if } \lambda < -1/2 \end{cases} \tag{B.35}$$

*Reference.* Barndorff-Nielsen and Blæsild (1981, pp. 38–39) ☐

**Remark B.36.** *Lorch (1967) proved that $R_\lambda$ is decreasing if and only if $\lambda > -1/2$ and increasing if and only if $\lambda < -1/2$.*

# Appendix C

# Fourth Parametrization of GH Distributions

For the implementation of GH pricing models in real world applications we recomment the use of the 4th parametrization. Then one can restrict to the parameters $\lambda, \overline{\alpha}, \overline{\beta}, \delta, \mu$ or $\lambda, \overline{\alpha}, \overline{\beta}, \sigma, \mu$, the volatility $\sigma$ denotes the standard deviation of the distribution. To compute a density or a pricing formula, one simply has to calculate $\delta$ with formula (2.1). The advantage is, that for a user, who is not familiar with details of GH distributions a variance parameter and some shape parameters are less confusing than the rescaling procedure. Moreover, from the representation of the densities in the fourth parametrization the role of the scale and location parameters $\delta$ resp. $\mu$ is obvious: in all formulae only the term $(\frac{x-\mu}{\delta})$ and not $x$ alone appears.

**Definition C.1.** *The generalized hyperbolic* (GH) *distribution is in the univariate case given by the following Lebesgue density*

$$
\begin{aligned}
\mathrm{gh}(x; \lambda, \overline{\alpha}, \overline{\beta}, \delta, \mu) \;=\;& a(\lambda, \overline{\alpha}, \overline{\beta}, \delta) \left(1 + (\tfrac{x-\mu}{\delta})^2\right)^{(\lambda - \frac{1}{2})/2} \\
& \times \mathrm{K}_{\lambda - 1/2}\left(\overline{\alpha}\sqrt{1 + (\tfrac{x-\mu}{\delta})^2}\right) \exp\left(\overline{\beta}(\tfrac{x-\mu}{\delta})\right)
\end{aligned}
\tag{C.2}
$$

$$
a(\lambda, \overline{\alpha}, \overline{\beta}, \delta) \;=\; \frac{(\overline{\alpha}^2 - \overline{\beta}^2)^{\lambda/2}}{\sqrt{2\pi}\,\overline{\alpha}^{\lambda - 1/2}\,\delta\,\mathrm{K}_\lambda\left(\sqrt{\overline{\alpha}^2 - \overline{\beta}^2}\right)}
\tag{C.3}
$$

*where* $\mathrm{K}_\lambda$ *is a modified Bessel function and* $x \in \mathbb{R}$. *The domain of variation of the parameters is* $\mu \in \mathbb{R}$ *and*

$$
\begin{aligned}
\delta \geq 0, |\overline{\beta}| < \overline{\alpha} \qquad & \text{if } \lambda > 0 \\
\delta > 0, |\overline{\beta}| < \overline{\alpha} \qquad & \text{if } \lambda = 0 \\
\delta > 0, |\overline{\beta}| \leq \overline{\alpha} \qquad & \text{if } \lambda < 0.
\end{aligned}
$$

**Definition C.4.** *For* $\lambda = 1$ *we obtain the* hyperbolic *distribution* (HYP)

$$
\frac{\sqrt{\overline{\alpha}^2 - \overline{\beta}^2}}{2\delta\overline{\alpha}\,\mathrm{K}_1(\sqrt{\overline{\alpha}^2 - \overline{\beta}^2})} \exp\left(-\overline{\alpha}\sqrt{1 + (\tfrac{x-\mu}{\delta})^2} + \overline{\beta}(\tfrac{x-\mu}{\delta})\right)
$$

*where* $x, \mu \in \mathbb{R}$, $0 \leq \delta$ *and* $0 \leq |\overline{\beta}| < \overline{\alpha}$.

**Definition C.5.** *For $\lambda = -1/2$ we get the* normal inverse Gaussian (NIG) *distribution:*

$$\frac{\bar{\alpha}}{\delta\pi} \exp\left(\sqrt{\bar{\alpha}^2 - \bar{\beta}^2} + \bar{\beta}(\tfrac{x-\mu}{\delta})\right) \frac{\mathrm{K}_1\left(\bar{\alpha}\sqrt{1 + (\tfrac{x-\mu}{\delta})^2}\right)}{\sqrt{1 + (\tfrac{x-\mu}{\delta})^2}}$$

$x, \mu \in \mathbb{R}$, $0 \leq \delta$ *and* $0 \leq |\bar{\beta}| \leq \bar{\alpha}$.

**Proposition C.6.** *The generalized hyperbolic distribution has the following mean and variance*

$$EX \quad = \quad \mu + \frac{\delta\bar{\beta}}{\zeta}\frac{\mathrm{K}_{\lambda+1}(\zeta)}{\mathrm{K}_\lambda(\zeta)} \tag{C.7}$$

$$VX \quad = \quad \delta^2 \left(\frac{\mathrm{K}_{\lambda+1}(\zeta)}{\zeta\,\mathrm{K}_\lambda(\zeta)} + \frac{\bar{\beta}^2}{\bar{\alpha}^2 - \bar{\beta}^2}\left[\frac{\mathrm{K}_{\lambda+2}(\zeta)}{\mathrm{K}_\lambda(\zeta)} - \left(\frac{\mathrm{K}_{\lambda+1}(\zeta)}{\mathrm{K}_\lambda(\zeta)}\right)^2\right]\right) \tag{C.8}$$

*where $\zeta = \sqrt{\bar{\alpha}^2 - \bar{\beta}^2}$. In (C.8) the term in the round brackets is scale- and location-invariant.*

# References

Abramowitz, M. and I. A. Stegun (1968). *Handbook of Mathematical Functions*. New York: Dover Publ.

Anderson, T. W. (1958). *An introduction to multivariate statistical analysis*. New York: Wiley.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1997). Thinking coherently. *Risk 10*, 68–71.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*, 203–228.

Atkinson, A. C. (1982). The simulation of generalized inverse Gaussian and hyperbolic random variables. *SIAM Journal on Scientific Computing 3*, 502–515.

Bakshi, G., C. Cao, and Z. Chen (1997). Empirical performance of alternative option pricing models. *Journal of Finance 52*, 2003–2049.

Bakshi, G., C. Cao, and Z. Chen (1998). Do call prices and the underlying stock move in opposite directions? Working paper, University of Maryland. To appear in Review of Financial Studies.

Barndorff-Nielsen, O. E. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society London A 353*, 401–419.

Barndorff-Nielsen, O. E. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of Statistics 5*, 151–157.

Barndorff-Nielsen, O. E. (1988). *Parametric Statistical Models and Likelihood*, Volume 50 of *Lecture Notes in Statistics*. Heidelberg: Springer.

Barndorff-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics 24*, 1–13.

Barndorff-Nielsen, O. E. (1998). Processes of normal inverse Gaussian type. *Finance & Stochastics 2*, 41–68.

Barndorff-Nielsen, O. E. (1999). Superposition of Ornstein-Uhlenbeck type processes. Ma-PhySto Research Report 2, University of Århus.

Barndorff-Nielsen, O. E. and P. Blæsild (1981). Hyperbolic distributions and ramifications: Contributions to theory and application. In C. Taillie, G. Patil, and B. Baldessari (Eds.), *Statistical Distributions in Scientific Work*, Volume 4, pp. 19–44. Dordrecht: Reidel.

Barndorff-Nielsen, O. E. and P. Blæsild (1983). Hyperbolic distributions. In S. Kotz and N. L. Johnson (Eds.), *Encyclopedia of Statistics*, Volume 3. New York: Wiley.

Barndorff-Nielsen, O. E., P. Blæsild, J. L. Jensen, and M. Sørensen (1985). The fascination of sand. In A. C. Atkinson and S. E. Fienberg (Eds.), *A Celebration of Statistics*. New York: Springer.

Barndorff-Nielsen, O. E. and O. Halgreen (1977). Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 38*, 309–312.

Barndorff-Nielsen, O. E., J. L. Jensen, and M. K. Sørensen (1995). Some stationary processes in discrete and continuous time. Research Report 241, Department of Theoretical Statistics, University of Århus.

Barndorff-Nielsen, O. E. and K. Prause (1999). Apparent scaling. Research Report 408, Department of Theoretical Statistics, University of Århus.

Barndorff-Nielsen, O. E. and N. Shephard (1998). Aggregation and model construction for volatility models. Working Paper 10, Centre for Analytical Finance, University of Århus.

Barndorff-Nielsen, O. E. and N. Shephard (1999). Non-Gaussian OU based models and some of their uses in financial economics. Working Paper 37, Centre for Analytical Finance, University of Århus.

Basel Committee on Banking Supervision (1995). *An internal model-based approach to market risk capital requirements*. Basel: Bank for International Settlements.

Basel Committee on Banking Supervision (1996). *Supervisory framework for the use of "backtesting" in conjunction with the internal models approach to market risk capital requirements*. Basel: Bank for International Settlements.

Basel Committee on Banking Supervision (1999). *A new capital adequacy framework*. Basel: Bank for International Settlements.

Bates, D. S. (1991). The crash of '87: was it expected? The evidence from options markets. *Journal of Finance 46*, 1009–1044.

Bates, D. S. (1996). Testing option pricing models. In G. S. Maddala and C. R. Rao (Eds.), *Handbook of Statistics*, Volume 14, pp. 567–611. Amsterdam: Elsevier Science.

Bates, D. S. (1997). Post-'87 crash fears in S&P 500 futures options. Working paper 5894, National Bureau of Economic Research.

Batún, J. L. and K. Prause (1999). Simulation of stationary and selfsimilar processes driven by Lévy processes. Working Paper, University of Århus.

Beckers, S. (1981). Standard deviations implied in option prices as predictors of future stock price variability. *Journal of Banking and Finance 5*, 661–673.

Beißer, J. (1999). Another way to value basket options. Working paper, Johannes Gutenberg-Universität Mainz.

Bibby, B. M. and M. Sørensen (1997). A hyperbolic diffusion model for stock prices. *Finance & Stochastics 1*, 25–41.

Blæsild, P. (1981). The two-dimensional hyperbolic distribution and related distributions, with an application to Johannsen's bean data. *Biometrika 68*, 251–263.

Blæsild, P. (1999). Generalized hyperbolic and generalized inverse Gaussian distributions. Working Paper, University of Århus.

Blæsild, P. and J. L. Jensen (1981). Multivariate distributions of hyperbolic type. In C. Taillie, G. Patil, and B. Baldessari (Eds.), *Statistical distributions in scientific work*, Volume 4, pp. 45–66. Dordrecht: Reidel.

Blæsild, P. and M. K. Sørensen (1992). 'hyp' – a computer program for analyzing data by means of the hyperbolic distribution. Research Report 248, Department of Theoretical Statistics, University of Århus.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics 69*, 542–547.

Breckling, J., E. Eberlein, and P. Kokic (1999). A new framework for the evaluation of market and credit risk.

Bühler, W., O. Korn, and A. Schmidt (1998). Ermittlung von Eigenkapitalanforderungen mit internen Modellen. Eine empirische Studie zur Messung von Zins-, Währungs- und Optionsrisiken mit Value-at-Risk Ansätzen. *Die Betriebswirtschaft 58*(1), 64–85.

Bühlmann, H., F. Delbaen, P. Embrechts, and A. N. Shiryaev (1996). No–arbitrage, change of measure and conditional Esscher transforms. *CWI Quarterly 9*, 291–317.

Chan, T. (1999). Pricing contingent claims on stocks driven by Lévy processes. *Annals of Applied Probability 9*, 504–528.

Christensen, B. J. and N. M. Kiefer (1999). Simulated moment methods for empirical equivalent martingale measures. Working Paper 31, Centre for Analytical Finance, University of Århus.

Christensen, B. J. and N. R. Prabhala (1998). The relation between implied and realized volatility. Working Paper 14, Centre for Analytical Finance, University of Århus.

Christiansen, C. (1999). Value-at-Risk using the factor-ARCH model. *Journal of Risk 1*, 65–86.

Cleveland, W. S., E. Grosse, and W. M. Shyu (1993). Local regression models. In J. M. Chambers and T. J. Hastie (Eds.), *Statistical Models in S*. London: Chapman & Hall.

Clewlow, L. and X. Xu (1993). The dynamics of stochastic volatility. FORC Preprint 94/53, University of Warwick.

Cont, R. (1999). Econometrics without probability: measuring the pathwise regularity of price trajectories. Talk at the workshop on *Product Integrals and Pathwise Integration*, January 11–13, 1999, MaPhySto, University of Århus.

Cormack, R. M. (1971). A review of classification. *Journal of the Royal Statistical Society A 134*, 321–367.

Cox, J. C. and M. Rubinstein (1985). *Options Markets*. Englewood Cliffs: Prentice Hall.

Cuppens, R. (1975). *Decomposition of multivariate probabilities*, Volume 29 of *Probability and Mathematical Statistics*. New York: Academic Press.

Daniels, H. E. (1987). Tail probability approximations. *International Statistical Review 55*, 37–48.

Davé, D. R. and G. Stahl (1997). On the accuracy of VaR estimates based on the variance-covariance approach. Working paper, Olsen & Associates.

Delbaen, F. and W. Schachermayer (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen 300*, 463–520.

Derman, E. and I. Kani (1994). The volatility smile and its implied tree. Quantitative strategies research notes, Goldman & Sachs.

Deutsche Börse (1997). Leitfaden zu den Volatilitätsindices der Deutschen Börse. www.exchange.de. Version 1.2.

Eberlein, E. (1998). Grundideen moderner Finanzmathematik. *Mitteilungen der DMV 3*, 10–20.

Eberlein, E., A. Ehret, O. Lübke, F. Özkan, K. Prause, S. Raible, R. Wirth, and M. Wiesendorfer Zahn (1998). *Freiburg Financial Data Tools*. Freiburg: Mathematische Stochastik, Universität Freiburg.

Eberlein, E. and J. Jacod (1997). On the range of options prices. *Finance and Stochastics 1*, 131–140.

Eberlein, E. and U. Keller (1995). Hyperbolic distributions in finance. *Bernoulli 1*, 281–299.

Eberlein, E., U. Keller, and K. Prause (1998). New insights into smile, mispricing and value at risk: the hyperbolic model. *Journal of Business 71*, 371–405.

Eberlein, E. and K. Prause (1998). The generalized hyperbolic model: financial derivatives and risk measures. FDM Preprint 56, University of Freiburg.

Eberlein, E. and S. Raible (1999). Term structure models driven by general Lévy processes. *Mathematical Finance 9*, 31–53.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997). *Extremal Events in Finance and Insurance*. Heidelberg: Springer.

Epps, T. W. (1979). Comovement in stock prices in the very short run. *Journal of the American Statistical Association 74*, 291–298.

Esscher, F. (1932). On the probability function in the collective theory of risk. *Skandinavisk Aktuarietidskrift 15*, 175–95.

Feller, W. (1966). *An introduction to probability theory and its applications* (2 ed.), Volume II. New York: Wiley.

Föllmer, H. and M. Schweizer (1991). Hedging of contingent claims under incomplete information. In M. H. A. Davis and R. J. Elliott (Eds.), *Applied Stochastic Analysis*, Volume 5 of *Stochastics Monographs*, pp. 389–414. New York: Gordon and Breach.

Föllmer, H. and D. Sondermann (1986). Hedging of non-redundant contingent claims. In W. Hildenbrand and A. Mas-Colell (Eds.), *Contributions to Mathematical Economics*, pp. 205–223. Amsterdam: North-Holland.

Frey, R. (1996). Derivative asset analysis in models with level dependent and stochastic volatility. *CWI Quaterly 10*, 1–34.

Geman, H., D. B. Madan, and M. Yor (1998). Asset prices are brownian motion: only in business time. Working Paper, University Paris IX Dauphine.

Gerber, H. U. and E. S. W. Shiu (1994). Option pricing by Esscher-transforms. *Transactions of the Society of Actuaries 46*, 99–191. With discussion.

Gilks, W. R. (1994). Markov chain Monte Carlo – a guide for practice. Working paper, Medical Research Council Biostatistics Unit, Cambridge, UK.

Grandits, P. (1996). The $p$-optimal martingale measure and its asymptotic relation with the Esscher transform. Preprint, Institut für Statistik, Universität Wien.

Grosswald, E. (1976). The Student t-distribution of any degree of freedom is infinitely divisible. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 36*, 103–109.

Guillaume, D. M., M. M. Dacorogna, R. D. Davé, U. A. Müller, R. B. Olsen, and O. V. Pictet (1997). From the bird's eye to the microscope: a survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics 1*, 95–129.

Gut, A. (1995). *An Intermediate Course in Probability*. New York: Springer.

Halgreen, C. (1979). Self-decomposability of the generalized inverse Gaussian and hyperbolic distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 47*, 13–17.

Huff, B. W. (1969). The strict subordination of differential processes. *Sankhya A31*, 403–412.

Huisman, R., K. G. Koedijk, and A. J. Pownall (1998). VaR-x: Fat tails in financial risk management. *Journal of Risk 1*, 47–61.

Hull, J. and A. White (1998). Incorporating volatility up-dating into the historical simulation method for value-at-risk. *Journal of Risk 1*, 5–19.

Hull, J. C. and A. White (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance 42*, 281–300.

Hurst, S. R., E. Platen, and S. T. Rachev (1995). Option pricing for asset returns driven by subordinated processes. Working paper, The Australian National University.

Iacus, S. M. and E. Nicolato (1999). A note on the existence of an equivalent martingale measure for a new class of stochastic volatility models. Working Paper, University of Padova.

Jarrat, A. (1970). A review of methods for solving nonlinear algebraic equations in one variable. In P. Rabinowitz (Ed.), *Numerical Methods for Nonlinear Algebraic Equations*. London: Gordon and Breach.

Jaschke, S. R. (1997). A note on stochastic volatility, GARCH models, and hyperbolic distributions. Working paper, SFB 373, Humboldt–Universität Berlin.

Jensen, J. L. (1995). *Saddlepoint Approximations*. Oxford: Clarendon Press.

Jiang, W. and J. Pedersen (1998). Parameter estimation for a discretely observed stochastic volatility model with jumps in the volatility. Working Paper, University of Århus.

Johnson, N. L. and S. Kotz (1970). *Continuous univariate distributions*, Volume 2. Boston: Houghton Mifflin.

Jørgensen, B. (1982). *Statistical properties of the generalized inverse Gaussian distribution*, Volume 9 of *Lecture Notes in Statistics*. Heidelberg: Springer.

Jorion, P. (1997). *Value at Risk. The new benchmark for controlling derivatives risk*. New York: McGraw-Hill.

J.P. Morgan and Reuters (1996). RiskMetrics – Technical document. New York.

Jurek, Z. J. and W. Vervaat (1983). An integral representation for selfdecomposable Banach space valued random variables. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 62*, 247–262.

Kallsen, J. and M. S. Taqqu (1998). Option pricing in ARCH–type models. *Mathematical Finance 8*, 13–26.

Keller, U. (1997). *Realistic modelling of financial derivatives*. Dissertation, University of Freiburg.

Kolb, R. (1995). *Understanding Options*. New York: Wiley.

Kotz, S. and N. L. Johnson (1983). *Encyclopedia of statistical sciences*, Volume 3. New York: John Wiley & Sons.

Lamberton, D. and B. Lapeyre (1996). *Introduction to Stochastic Calculus applied to Finance*. London: Chapman & Hall.

Liptser, R. S. and A. N. Shiryayev (1977). *Statistics of Random Processes I. General Theory*, Volume 5 of *Applications of Mathematics*. New York: Springer.

Lo, A. W. (1986). Statistical tests of contingent-claims asset-pricing models. A new methodology. *Journal of Financial Economics 17*, 143–173.

Longstaff, F. A. (1995). Stochastic volatility and option valuation: a pricing-density approach. Working paper, University of California, Los Angeles.

Lopez, J. A. (1999). Regulatory evaluation of value-at-risk models. *Journal of Risk 1*, 37–63.

Lorch, L. (1967). Inequalities for some Whittaker functions. *Arch. Math. (Brno) 3*, 1–9.

Loretan, M. (1997). Generating market risk scenarios using principal components analysis: methodological and practial considerations. In Bank for International Settlement (Ed.), *The Measurement of Aggregate Market Risk*, pp. 23–60. Basel: Bank for International Settlement.

Lugannani, R. and S. Rice (1980). Saddle point approximation for the distribution of the sum of independent random variables. *Adv. Appl. Prob. 12*, 475–490.

Lukasz, E. (1970). *Characteristic functions*. London: Griffin.

Madan, D. B., P. P. Carr, and E. C. Chang (1998). The variance gamma process and option pricing. *European Finance Review 2*, 79–105.

Mandelbrot, B. B. (1963). The variation of certain speculative prices. *Journal of Business 36*, 394–419.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika 57*, 519–530.

Marinelli, C., S. Rachev, and R. Roll (1999). Subordinated exchange rate models: Evidence for heavy tailed distributions and long-range dependence. To appear in Stable Models in Finance.

Marinelli, C., S. Rachev, R. Roll, and H. Göppl (1999). Subordinated stock price models: Heavy tails and long-range dependence in the high-frequency Deutsche Bank price record. Working Paper.

Matthes, R. and M. Schröder (1998). Portfolio analysis based on the shortfall concept. In G. Bol, G. Nakhaeizadeh, and K.-H. Vollmer (Eds.), *Risk Measurement, Econometrics and Neural Networks*, Heidelberg, pp. 147–160. Physica-Verlag. Selected Articles of the 6th Econometric-Workshop in Karlsruhe, March 19–21, 1997.

Michael, J. R., W. R. Schucany, and R. W. Haas (1979). Generating random variables using transformations with roots. *The American Statistician 30*, 88–89.

Musiela, M. and M. Rutkowski (1997). *Martingale Methods in Financial Modelling*, Volume 36 of *Applications of Mathematics*. Heidelberg: Springer.

Nelson, D. B. (1990). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory 6*, 318–334.

Nicolato, E. and K. Prause (1999). Derivative pricing in stochastic volatility models of the Ornstein-Uhlenbeck type. Working Paper, University of Århus.

Olsen & Associates (1998). *Second International Conference on High Frequency Data in Finance, April 1–3, 1998*, Zürich. Olsen & Associates.

Özkan, F. (1997). Intraday-Modelle für Aktienkurse. Diplomarbeit, Institut für Mathematische Stochastik, Universität Freiburg.

Pfanzagl, J. (1994). *Parametric Statistical Theory*. Berlin: de Gruyter.

Prause, K. (1997). Modelling financial data using generalized hyperbolic distributions. FDM Preprint 48, University of Freiburg.

Prause, K. (1999a). The generalized hyperbolic model: financial derivatives and risk measures. In *Eleventh European Young Statisticians Meeting, Marly-le-Roi, August 24–28, 1999*, pp. 165–169. Bernoulli Society.

Prause, K. (1999b). How to use NIG laws to measure market risk. FDM Preprint 65, University of Freiburg.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery (1992). *Numerical Recipes in C*. Cambridge: Cambridge University Press.

Raible, S. (1998). Lévy term-structure models: Uniqueness of the martingale measure. Working paper, University of Freiburg.

Raible, S. (1999). Dissertation, University of Freiburg. To appear.

Rawnsley, J. (1995). *Total risk: Nick Leeson and the Fall of Barings Bank*. New York: Harper Business.

Revuz, D. and M. Yor (1999). *Continuous martingales and Brownian motion* (3 ed.), Volume 293 of *Grundlehren der mathematischen Wissenschaften*. Berlin: Springer.

Roberts, G. O. and A. F. M. Smith (1994). Simple conditions for the convergence of the Gibbs sampler and Metropolis–Hastings algorithms. *Stochastic Processes and their Applications 49*, 207–216.

Rodrigues, P. (1997). Term structure and volatility shocks. In Bank for International Settlement (Ed.), *The measurement of aggregate market risk*, pp. 23–60. Basel: Bank for International Settlement.

Rogers, L. C. G. and O. Zane (1999). Saddle-point approximations to option prices. *Annals of Applied Probability 9*, 493–503.

Rosinski, J. (1991). On a class of infinitely divisible processes represented as mixtures of Gaussian processes. In S. Cambanis, G. Samorodnitzky, and M. S. Taqqu (Eds.), *Stable processes and related topics*, pp. 27–41. Basel: Birkhäuser.

Rubinstein, M. (1985). Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978. *Journal of Finance 40*, 455–480.

Rydberg, T. H. (1997a). The normal inverse Gaussian Lévy process: Simulation and approximation. *Communications in Statistics: Stochastic Models 13*, 887–910.

Rydberg, T. H. (1997b). A note on the existence of unique equivalent martingale measures in a Markovian setting. *Finance and Stochastics 1*, 251–257.

Rydberg, T. H. (1998). *Some Modelling Results in the Area of Interplay between Statistics, Mathematical Finance, Insurance and Econometrics*. PhD.-thesis, Memoir 14, Department of Theoretical Statistics, University of Århus.

Rydberg, T. H. (1999). Generalized hyperbolic diffusion processes with applications towards finance. *Mathematical Finance 9*, 183–201.

Samuelson, P. (1965). Rational theory of warrant pricing. *Industrial Management Review 6*, 13–32.

Sato, K., T. Watanabe, and M. Yamazato (1994). Recurrence conditions for multidimensional processes of Ornstein–Uhlenbeck type. *J. Math. Soc. Japan 46*, 245–265.

Sato, K. and M. Yamazato (1982). Stationary processes of Ornstein-Uhlenbeck type. In K. Itô and J. V. Prohorov (Eds.), *Probability Theory and Mathematical Statistics*, Volume 1021 of *Lecture Notes in Mathematics*. Berlin: Springer. Proceedings of the 4th USSR–Japan Symposium.

Schnidrig, R. and D. Würtz (1995). Investigation of the volatility and autocorrelation function of the USDDEM exchange rate on operational time scales. In *Proceedings of the "1st International Conference on High Frequency Data"*, Zürich. March 29–31, 1995.

Schweizer, M. (1988). *Hedging of Options in a General Semimartingale Modell*. Dissertation no. 8615, ETHZ.

Schweizer, M. (1991). Option hedging for semimartingales. *Stochastic Processes and their Applications 37*, 339–363.

Schweizer, M. (1992). Mean-variance hedging for general claims. *The annals of applied probability 2*, 171–179.

Seber, G. A. F. (1984). *Multivariate Observations*. New York: John Wiley & Sons.

Shiryaev, A. (1999). *Essentials of stochastic finance: facts, models, theory*. Singapore: World Scientific.

Soni, R. P. (1965). On an inequality for modified Bessel functions. *Jour. Math. and Physics 44*, 406–407.

Sulem, A. (1999). Optimal portfolio and concumption in a jump diffusion market. Talk at the *Conference on Lévy processes: theory and applications*, January 18–22, 1999. MaPhySto Miscellanea 11, University of Århus.

Thisted, R. A. (1988). *Elements of Statistical Computing*. New York, London: Chapman & Hall.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *Annals of Statistics 22*, 1701–1728.

Watson, G. N. (1952). *A treatise on the theory of Bessel functions* (2 ed.). Cambridge: University Press.

Wiesendorfer Zahn, M. (1999). Simulation von Lévy-Prozessen. Diplomarbeit, Institut für Mathematische Stochastik, Universität Freiburg.

Wirth, R. (1998). Stochastische Modellierung von Volatilitätsprozessen. Diplomarbeit, Institut für Mathematische Stochastik, Universität Freiburg.

Witting, H. (1985). *Mathematische Statistik I. Parametrische Verfahren bei festem Stichprobenumfang.* Stuttgart: Teubner.

Witting, H. and U. Müller-Funk (1995). *Mathematische Statistik II. Asymptotische Statistik: Parametrische Modelle und nichtparametrische Funktionale.* Stuttgart: Teubner.

Wolfe, S. J. (1982). On a continuous analogue of the stochastic difference equation $X_n = \varrho X_{n+1} + B_n$. *Stoch. Prob. Appl. 12*, 301–312.