

Chen, Ying; Härdle, Wolfgang Karl; Jeong, Seok-Oh

**Working Paper**

## Nonparametric risk management with generalized hyperbolic distributions

SFB 649 Discussion Paper, No. 2005-001

**Provided in Cooperation with:**

Collaborative Research Center 649: Economic Risk, Humboldt University Berlin

*Suggested Citation:* Chen, Ying; Härdle, Wolfgang Karl; Jeong, Seok-Oh (2005) : Nonparametric risk management with generalized hyperbolic distributions, SFB 649 Discussion Paper, No. 2005-001, Humboldt University of Berlin, Collaborative Research Center 649 - Economic Risk, Berlin

This Version is available at:

<https://hdl.handle.net/10419/25020>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Nonparametric Risk Management with Generalized Hyperbolic Distributions

Ying Chen\*  
Wolfgang Härdle\*  
Seok-Oh Jeong\*\*



\* CASE - Center for Applied Statistics and Economics,  
Humboldt-Universität zu Berlin, Germany

\*\* Institute de Statistique, Université Catholique de Louvain,  
Belgium

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>  
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin  
Spandauer Straße 1, D-10178 Berlin



# Nonparametric Risk Management with Generalized Hyperbolic Distributions

Chen, Ying\*    Härdle, Wolfgang\*    and    Jeong, Seok-Oh\*\*

\*CASE - Center for Applied Statistics and Economics

Humboldt-Universität zu Berlin

Wirtschaftswissenschaftliche Fakultät

Spandauerstrasse 1, 10178 Berlin

Germany

\*\* Institut de Statistique

Université Catholique de Louvain

Voie du Roman Pays, 20 1348 Louvain-la-Neuve

Belgium

1st August 2005

## Abstract

In this paper we propose the GHADA risk management model that is based on the generalized hyperbolic (GH) distribution and on a nonparametric adaptive methodology. Compared to the normal distribution, the GH distribution possesses semi-heavy tails and represents the financial risk factors more appropriately. The nonparametric adaptive methodology has the desirable property of estimating homogeneous volatility in a short time interval. For DEM/USD exchange rate data and a German bank portfolio data the proposed GHADA model provides more accurate value at risk calculation than the traditional model based on the normal distribution. All calculations and simulations are done with XploRe.

*Keywords:* adaptive volatility estimation, generalized hyperbolic distribution, value at risk, risk management.

*Acknowledgement:* This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 “Economic Risk” and the SFB 373 “Simulation and Quantification of

Economic Processes” at Humboldt-Universität zu Berlin. Special thanks are due to Prof. Dr. Ernst Eberlein for his kind contribution to the proof of Lemma 1.

## 1 INTRODUCTION

One of the most challenging tasks in the analysis of financial markets is to measure and manage risks properly. After the breakdown of the fixed exchange rate system of the Bretton Woods Agreement in 1971, a sudden increase of volatility was observed in the financial markets. The following boom of financial derivatives accelerated the turbulence of the markets. The incoming scale of losses astonished the world and pushed the development of sound risk management systems. Financial risks have many sources and are typically mapped into a stochastic framework where various kinds of risk measures such as Value at Risk (VaR), expected shortfall, lower partial moments are calculated. Among them, VaR has become the standard measure of the market risk since J.P. Morgan launched RiskMetrics in 1994, making the analysis of VaR simple and standard, Jorion (2001).

The importance of VaR was even reinforced after it was used by the central banks to govern and supervise the capital adequacy of the banks in the Group of Ten (G10) countries in 1995. Mathematically VaR at  $p$  probability level is defined as:

$$VaR_{p,t} = F_t^{-1}(p),$$

where  $F_t^{-1}$  is the inverse function of the conditional cumulative distribution function of the underlying at time  $t$ , Franke, Härdle and Hafner (2004). From the definition, it is clear that the accuracy of VaR and the other risk measures depends heavily on the assumption of the underlying distribution. In the literature, for reasons of stochastic and numerical simplicity, it is often assumed that the involved risk factors are normally distributed. This is done e.g. in the RiskMetrics framework. However empirical studies have shown that financial risk factors have leptokurtic distributions which include a high peak and fat tails.

Figure 1 illustrates this fact on the basis of the daily standardized (devolatilized) returns of the foreign exchange (FX) rates DEM/USD from 1979/12/01 to 1994/04/01. The non-parametrically estimated kernel density and log kernel density obviously deviate from the normal density. In order to capture this empirical fact, the heavy-tailed distribution families such as the hyperbolic distribution have been attracting the attention of the researchers. Conditional Gaussian models can mimic the fat tails as well and were found to perform well at a moderate VaR (e.g. 95%) confidence level. Nevertheless, they are unsatisfactory for the extreme events such as profit and loss (P&L) at 99% confidence level, Jaschke and Jiang (2002). Recently, Eberlein, Kallsen and Kristen (2003) applied the generalized hyperbolic

(GH) distribution to the VaR calculation. Based on their empirical studies, the model with GH distribution gave more accurate VaR values than that with the normal distribution.

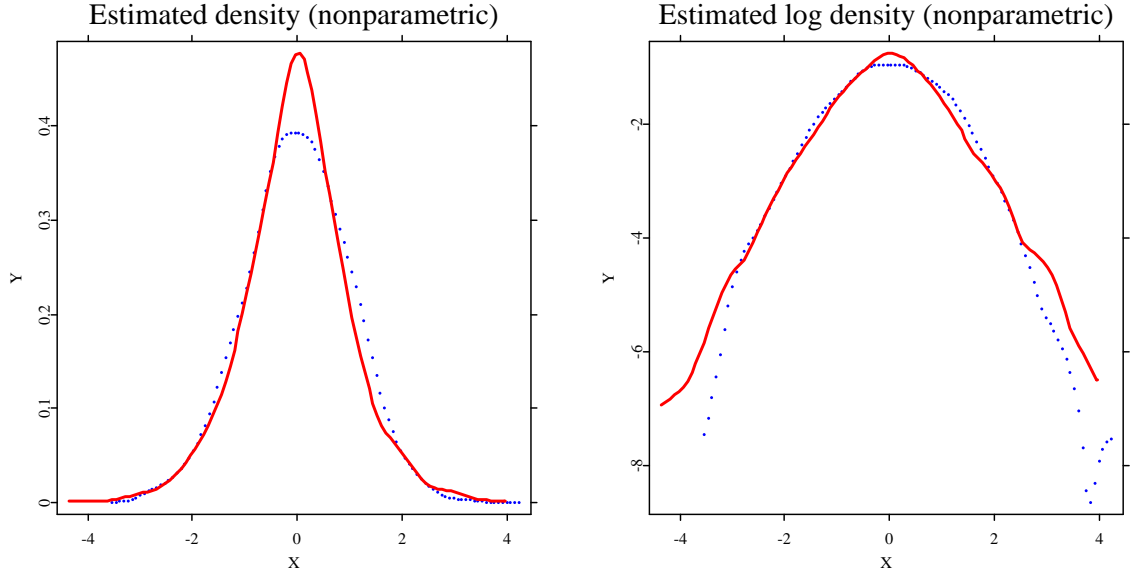


Figure 1: Graphical comparison of the density (left) and the log-density (right) of the daily DEM/USD standardized returns from 1979/12/01 to 1994/04/01 (3719 observations). The kernel density estimate is graphed as a line and the normal density as dots with  $h \approx 0.54$ .

[GHADafx.xpl](#)

In addition to the distributional tail assumption, the usual heteroscedastic model on the returns  $R_t$ :

$$R_t = \sigma_t \varepsilon_t,$$

where  $\sigma_t$  denotes the volatility and  $\varepsilon_t$  the stochastic term, suggests that the role of the volatility model is of great significance. The most often used volatility estimation methods are not uniformly applicable in risk management. ARCH (Engle, 1995), GARCH (Bollerslev, 1995) and stochastic volatility models (Harvey and Shephard, 1995) are used to estimate or forecast volatility in specified time periods. This cannot be applied for long time series since the form of the volatility model is time-unstable with a very high possibility. It is therefore plausible to use more flexible methods by providing a data-driven “local” model, which can avoid this potential miss-specification problem. In Eberlein et al. (2003), parametric volatility models of GARCH type were studied and compared with a nonparametric approach using the rectangular moving average. They argued that the GARCH (1,1) model performed superior. However this “under-performance” of the nonparametric model can possibly be improved by an adaptive methodology. Mercurio and

Spokoyny (2004) proposed such an improvement by adaptively estimating the volatility. A simple “local constant” model was constructed but an intrinsic assumption in their study was the normality of the risk factors.

In this paper we intend to improve the risk management models by combining:

- a. a heavy-tailed distribution family to mimic the empirical distribution of the underlying risk factors, and
- b. a nonparametric adaptive methodology to estimate and forecast the local volatilities.

Motivated by the above two research lines, we combine these two approaches: we estimate the volatility adaptively and model the heavy-tailed risk factors by the GH distribution. Here we name this new VaR technique as **Generalized Hyperbolic Adaptive Volatility (GHADA)** technique. The devolatilized return density plot in Figure 1 is in fact calculated with the GHADA technique.

The paper is organized as follows. In Section 2 we will discuss the properties of the GH distribution and its subclasses. The adaptive volatility estimation methodology will be described based on the GH distribution. The validation of the GHADA technique is shown via Monte Carlo simulation. In Section 3 several VaR calculations will be presented based on a DEM/USD series and a German bank portfolio data. According to the backtesting result, the GHADA technique provides more accurate precision than the model with the normal distribution. Finally we will conclude our study in Section 4. All the pictures may be recalculated and redrawn using the indicated link to an XploRe Quantlet Server.

## 2 PILLARS

Let  $R_t = \log S_t - \log S_{t-1}$  denote the (log) return where  $S_t$  is the asset price at time point  $t$  for  $t = 1, 2, \dots, T$ . The return process is modelled in a heteroscedastic form:

$$R_t = \sigma_t \varepsilon_t, \tag{1}$$

where  $\varepsilon_t$  is assumed to be independently and identically distributed (i.i.d.) with  $E(\varepsilon_t) = 0$  and  $\text{Var}(\varepsilon_t) = 1$ . Volatility  $\sigma_t$  is time varying and unobservable in the market. In the case that volatility is measurable with respect to a  $\sigma$ -field  $\mathcal{F}_{t-1}$  generated by the preceding returns  $R_1, \dots, R_{t-1}$ , the variance  $\sigma_t^2$  can be interpreted as the conditional variance of the return.

In risk management models we are interested in estimating the future return distribution accurately. It depends on the estimation of volatility  $\sigma_t$  and the assumption on the

stochastic term  $\varepsilon_t$ . In this paper, we use a nonparametric algorithm to estimate the volatility “locally”, avoiding the potential mis-specification problem. Details will be discussed in Section 2.2. Another key factor, the distribution assumption of the stochastic term, influences the performance of the risk management procedures to a great extent. VaR is defined through a pre-decided quantile of the P&L distribution. Models based on the normality assumption achieve almost the same values at the 5% quantile (95% confidence level) as those with a leptokurtic (more heavy tailed) distribution. This explains to a certain extent the popularity of the normal distribution in the risk management models although the financial risk factors are empirically leptokurtic distributed. However concerning the extreme events, we need to consider lower quantiles such as 1% quantile. The difference relative to the normal becomes larger for lower quantiles of course. In this case, the normality assumption becomes invalid. Therefore in the next section, we concentrate on a heavy-tailed distribution, the generalized hyperbolic distribution, and discuss the application of this distribution family in risk management.

## 2.1 Generalized Hyperbolic Distribution

The GH distribution introduced by Barndorff-Nielsen (1977) is a heavy-tailed distribution that can well replicate the empirical distribution of the financial risk factors. The density of the GH distribution for  $x \in \mathbb{R}$  is:

$$f_{GH}(x; \lambda, \alpha, \beta, \delta, \mu) = \frac{(\iota/\delta)^\lambda}{\sqrt{2\pi}K_\lambda(\delta\iota)} \frac{K_{\lambda-1/2}\left\{\alpha\sqrt{\delta^2 + (x-\mu)^2}\right\}}{\left\{\sqrt{\delta^2 + (x-\mu)^2}/\alpha\right\}^{1/2-\lambda}} \cdot e^{\beta(x-\mu)} \quad (2)$$

under the conditions:

- $\delta \geq 0, |\beta| < \alpha$  if  $\lambda > 0$
- $\delta > 0, |\beta| < \alpha$  if  $\lambda = 0$
- $\delta > 0, |\beta| \leq \alpha$  if  $\lambda < 0$

where  $\lambda, \alpha, \beta, \delta$  and  $\mu \in \mathbb{R}$  are the GH parameters,  $\iota^2 = \alpha^2 - \beta^2$ . The density's location and scale are mainly controlled by  $\mu$  and  $\delta$  respectively:

$$\begin{aligned} \mathbb{E}[X] &= \mu + \frac{\delta^2\beta}{\delta\iota} \frac{K_{\lambda+1}(\delta\iota)}{K_\lambda(\delta\iota)} \\ \text{Var}[X] &= \delta^2 \left\{ \frac{K_{\lambda+1}(\delta\iota)}{\delta\iota K_\lambda(\delta\iota)} + \left(\frac{\beta}{\iota}\right)^2 \left[ \frac{K_{\lambda+2}(\delta\iota)}{K_\lambda(\delta\iota)} - \left\{ \frac{K_{\lambda+1}(\delta\iota)}{K_\lambda(\delta\iota)} \right\}^2 \right] \right\}, \end{aligned}$$

whereas  $\beta$  and  $\alpha$  play roles in the skewness and kurtosis of the distribution. For more details of the parameters' domains we refer to Bibby and Sørensen (2001).  $K_\lambda(\cdot)$  is the modified

Bessel function of the third kind with index  $\lambda$ , Barndorff-Nielsen and Blæsild (1981):

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty y^{\lambda-1} \exp\left\{-\frac{x}{2}(y + y^{-1})\right\} dy$$

Furthermore, the GH distribution has semi-heavy tails:

$$f_{GH}(x; \lambda, \alpha, \beta, \delta, \mu = 0) \sim x^{\lambda-1} e^{-(\alpha-\beta)x} \text{ as } x \rightarrow \infty, \quad (3)$$

where  $a(x) \sim b(x)$  as  $x \rightarrow \infty$  means that both  $a(x)/b(x)$  and  $b(x)/a(x)$  are bounded as  $x \rightarrow \infty$ . Compared to the normal distribution, the GH distribution decays more slowly. However compared to the other two heavy-tailed distributions: Laplace distribution and Cauchy distribution, the decaying speed of the GH distribution is often faster. The Laplace distribution is also called double exponential distribution with the form:

$$f_{Laplace} = \frac{1}{2\varsigma} e^{-|x-\mu|/\varsigma}$$

where  $\mu$  is the location parameter and  $\varsigma$  is the scale parameter.

The Cauchy distribution is defined as:

$$f_{Cauchy} = \frac{1}{\varsigma \pi [1 + (x - M)^2/\varsigma^2]}$$

where  $M$  is the median and  $\varsigma$  is the scale parameter.

In Figure 2 we compared these four distributions and especially their tail-behavior. In order to keep the comparability of these distributions, we specified the means to 0 and standardized the variances to 1. Furthermore we used one important subclasses of the GH distribution: the normal-inverse Gaussian (NIG) distribution with  $\lambda = -\frac{1}{2}$  introduced more precisely in the following text. On the left panel, the complete forms of these distributions are revealed. The Cauchy (dots) distribution has the lowest peak and the fattest tails, in other words, it has the flattest distribution. The NIG distribution decays second fast in the tails although it has the highest peak, which is more clearly displayed on the right panel. Generally the GH distribution has an exponential decaying speed as shown in (3). By changing  $\lambda$  the GH distribution family covers a wide range of tail behavior. A parameter  $\lambda > 1$  introduces heavier tails than the double exponential. One may therefore think of  $\lambda$  as the tail control parameter with which one (loosely speaking) may model the range between a normal and a Cauchy tail.

The moment generating function of the GH distribution is:

$$m_f(z) = e^{\mu z} \cdot \frac{\iota^\lambda}{\iota_z^\lambda} \cdot \frac{K_\lambda(\delta \iota_z)}{K_\lambda(\delta \iota)}, \quad |\beta + z| < \alpha, \quad (4)$$



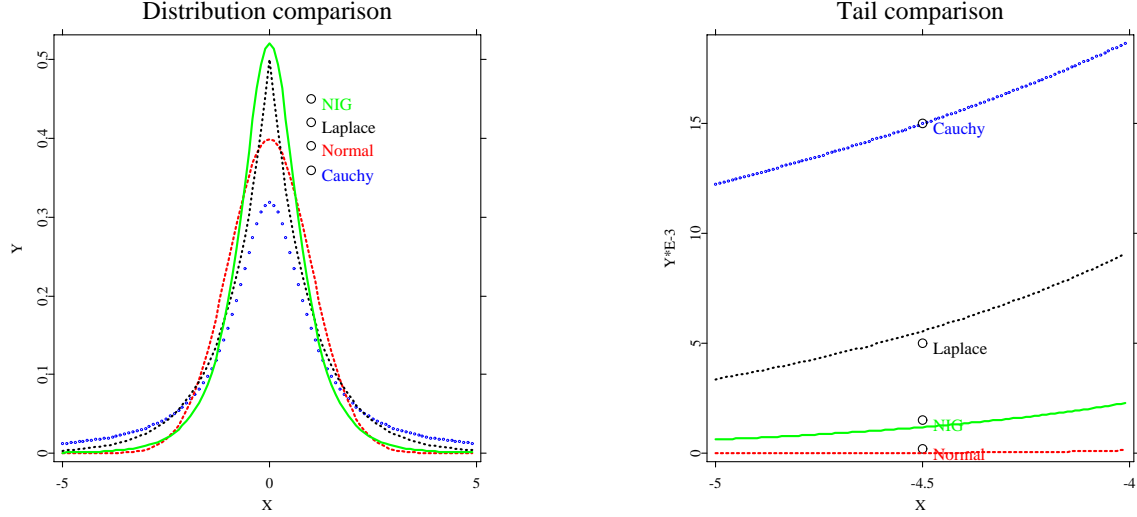


Figure 2: Graphical comparison of the NIG distribution (line), standard normal distribution (dashed), Laplace distribution (dotted) and Cauchy distribution (dots).

 [GHADAtail.xpl](#)

where  $\iota_z^2 = \alpha^2 - (\beta + z)^2$ . The GH distribution has the property that  $m_f$  is infinitely many times differentiable near 0, as a result every moment of a GH variable exists. In Section 2.2, this feature and the tail behavior (3) of the GH distribution will be used in the adaptive volatility estimation methodology.

In the current literature, subclasses of the GH distribution such as the hyperbolic (HYP) or the normal-inverse Gaussian (NIG) distribution are frequently used. This is motivated by the fact that the four parameters  $(\mu, \delta, \beta, \alpha)^\top$  simultaneously control the four moment functions of the distribution, i.e. the trend, the riskiness, the asymmetry and the likeliness of the extreme events. Eberlein and Keller (1995), Barndorff-Nielsen (1997) have shown that these subclasses are rich enough to model financial time series and have the benefit of numerical tractability. Therefore in our study we concentrate ourselves on these two important subclasses of the GH distribution: HYP with  $\lambda = 1$  and NIG distribution with  $\lambda = -1/2$ . The corresponding density functions are given as:

- Hyperbolic (HYP) distribution:  $\lambda = 1$ ,

$$f_{HYP}(x; \alpha, \beta, \delta, \mu) = \frac{\iota}{2\alpha\delta K_1(\delta\iota)} e^{\{-\alpha\sqrt{\delta^2 + (x-\mu)^2} + \beta(x-\mu)\}}, \quad (5)$$

where  $x, \mu \in \mathbb{R}$ ,  $0 \leq \delta$  and  $|\beta| < \alpha$ ,

- Normal-inverse Gaussian (NIG) distribution:  $\lambda = -1/2$ ,

$$f_{NIG}(x; \alpha, \beta, \delta, \mu) = \frac{\alpha\delta}{\pi} \frac{K_1 \left\{ \alpha \sqrt{\delta^2 + (x - \mu)^2} \right\}}{\sqrt{\delta^2 + (x - \mu)^2}} e^{\{\delta\alpha + \beta(x - \mu)\}}. \quad (6)$$

where  $x, \mu \in \mathbb{R}$ ,  $\delta > 0$  and  $|\beta| \leq \alpha$ .

In order to estimate the unknown parameters  $(\alpha, \beta, \delta, \mu)^\top$ , the maximum likelihood (ML) and numerical optimization methods are used. For an i.i.d HYP resp. NIG distributed variable  $X$ , the log-likelihood functions are:

$$L_{HYP} = T \log \iota - T \log 2 - T \log \alpha - T \log \delta - T \log K_1(\delta\iota) \quad (7)$$

$$+ \sum_{t=1}^T \{-\alpha \sqrt{\delta^2 + (x_t - \mu)^2} + \beta(x_t - \mu)\}$$

$$L_{NIG} = T \log \alpha + T \log \delta - T \log \pi + T \delta \iota \quad (8)$$

$$+ \sum_{t=1}^T \left[ \log K_1 \left\{ \alpha \sqrt{\delta^2 + (x_t - \mu)^2} \right\} - \frac{1}{2} \log \{\delta^2 + (x_t - \mu)^2\} + \beta(x_t - \mu) \right]$$

Figure 3 shows the estimated HYP and NIG densities with the corresponding ML estimators of the DEM/USD devolatilized returns. It can be seen that the estimated densities almost coincide with the empirical density and log-density of the financial risk factor. The empirical density  $\hat{f}_h(x)$  (line) was estimated by the kernel estimation:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right), \quad (9)$$

where  $n$  is the number of observations and  $K$  is the kernel function which gives weights to the observations according to the distances of them to the fixed point  $x$ . The further an observation is from the fixed point, the smaller weight it will be given. We chose the Quartic kernel function with a closed form:  $K(u) = \frac{15}{16}(1 - u^2)^2 1(|u| \leq 1)$ , where  $1(\cdot)$  is the indicator function which has a value of 1 if the condition in the parenthesis exists, and a value of 0 otherwise. In addition, we used the Silverman's rule of thumb to select the bandwidth  $h$ :

$$\hat{h}_{rot} \approx 1.06 \hat{\sigma} n^{-1/5},$$

where  $\hat{\sigma}$  is the empirical standard deviation of the variable  $X$ . Since the rule of thumb assumes that the unknown density belongs to the normal family and we chose the Quartic kernel, the bandwidth was adjusted to  $\hat{h} \approx 2.62 \hat{h}_{rot} = 2.78 \hat{\sigma} n^{-1/5}$  using the canonical bandwidths. For details of the kernel and the bandwidth selections, see Chapter 3 in Härdle, Müller, Sperlich and Werwatz (2004). Compared to the normal distribution in Figure 1, it is convincing that the GH distribution family can represent the empirical distribution of financial data better.

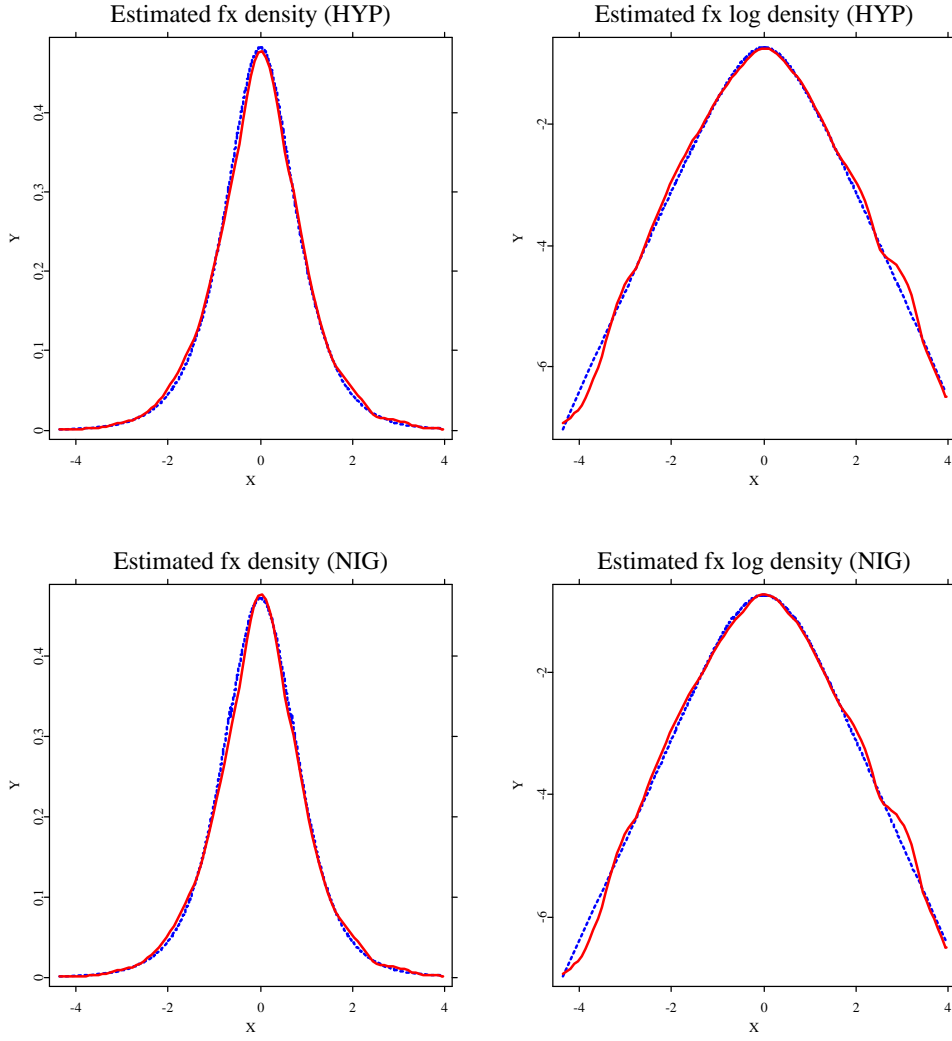


Figure 3: The kernel estimated density (left) and log density (right) of the standardized return of FX rates (line) ( $h \approx 0.55$ ). The HYP density (dashed line) on the top with the maximum likelihood estimators  $\hat{\alpha} = 1.744$ ,  $\hat{\beta} = -0.017$ ,  $\hat{\delta} = 0.782$ ,  $\hat{\mu} = 0.012$  and a simulated NIG density (dashed line) on the bottom with the maximum likelihood estimators  $\hat{\alpha} = 1.340$ ,  $\hat{\beta} = -0.015$ ,  $\hat{\delta} = 1.337$ ,  $\hat{\mu} = 0.010$ .

 [GHADafx.xpl](#)

## 2.2 Adaptive Volatility Estimation

The basic idea of adaptive volatility estimation comes from the observation that although the volatility is heteroscedastic in a long time period, its change in a short time interval, a so-called *time homogeneous interval*, is very small. Evidence for this argument has been given by Mercurio and Spokoiny (2004). According to time homogeneity, one specifies an interval  $I = [\tau - m, \tau)$  for a fixed time point  $\tau$  with  $0 \leq m \leq \tau - 1$ , where the volatility  $\sigma_t$ ,  $t \in I$  is almost constant. One may for example estimate in this case the local constant volatility  $\sigma_\tau$  by

averaging

the past squared returns  $R_t^2$  for  $t \in I$ :

$$\hat{\sigma}_\tau^2 = \frac{1}{|I|} \sum_{t \in I} R_t^2, \quad (10)$$

where  $|I|$  is the cardinality of  $I$ . Two questions arise in this procedure: how well does this estimate work and how to specify the homogeneous interval  $I$ ?

The squared returns  $R_t^2$  are always nonnegative and have for the stochastic errors  $\varepsilon_t$  (i.i.d. GH, HYP or NIG) a skewed distribution. In order to apply an interval selection procedure for  $I$  we use the power transformation of the return  $R_t$ :

$$|R_t|^\gamma = C_\gamma \sigma_t^\gamma + D_\gamma \sigma_t^\gamma \zeta_{\gamma,t} \quad (11)$$

where  $\gamma$  is the power transformation parameter,  $\zeta_{\gamma,t} = (|\varepsilon_t|^\gamma - C_\gamma)/D_\gamma$  are standardized i.i.d. innovations,  $C_\gamma = \mathbb{E}[|\varepsilon|^\gamma | \mathcal{F}_{t-1}]$  is the conditional mean of  $\zeta_{\gamma,t}$  and  $D_\gamma^2 = \mathbb{E}[(|\varepsilon|^\gamma - C_\gamma)^2 | \mathcal{F}_{t-1}]$  is the conditional variance of  $\zeta_{\gamma,t}$ .

Additionally, lighter tails are obtained after this power transformation. Such a tail behavior is required later to derive theoretical properties of the estimate later. Let us denote the conditional mean of the transformed return  $|R_t|^\gamma$  by  $\theta_t$ :

$$\theta_t = C_\gamma \sigma_t^\gamma. \quad (12)$$

Since  $C_\gamma$  is a constant given a fixed  $\gamma$ , the estimate of the volatility  $\sigma_t$  is proportional to  $\theta_t$ . In a time homogeneous interval  $I$  the local parameter  $\theta_t$  is constant for  $t \in I$  and is estimated by:

$$\hat{\theta}_I = \frac{1}{|I|} \sum_{t \in I} |R_t|^\gamma. \quad (13)$$

Writing  $|R_t|^\gamma$  out in (13), one has:

$$\hat{\theta}_I = \frac{1}{|I|} \sum_{t \in I} \theta_t + \frac{s_\gamma}{|I|} \sum_{t \in I} \theta_t \zeta_t$$

where  $s_\gamma = D_\gamma/C_\gamma$ . One sees that the multiplicative error structure (1) is turned via (11) into an additive one and the random variable  $|R_t|^\gamma$  distributes more evenly. Straightforwardly, one can calculate the conditional expectation and variance of the estimate  $\hat{\theta}_I$ :

$$\begin{aligned} \mathbb{E}[\hat{\theta}_I | \mathcal{F}_{\tau-1}] &= \mathbb{E} \left[ \frac{1}{|I|} \sum_{t \in I} \theta_t \right], \\ v_I^2 = \text{Var} [\hat{\theta}_I | \mathcal{F}_{\tau-1}] &= \frac{s_\gamma^2}{|I|^2} \mathbb{E} \left( \sum_{t \in I} \theta_t \zeta_t \right)^2 = \frac{s_\gamma^2}{|I|^2} \mathbb{E} \sum_{t \in I} \theta_t^2. \end{aligned}$$

In a time homogeneous interval  $I$ , the volatilities are expected to be time invariant, therefore  $\hat{\theta}_I$  can be considered as an estimate of  $\theta_t$  for each time point  $t \in I$ . Therefore  $v_I$  can be estimated by:

$$\hat{v}_I = s_\gamma \hat{\theta}_I |I|^{-1/2}.$$

In other words, the volatility estimate  $\sigma_t$  can be induced from an estimate  $\theta_t$ . However the specification of the local homogeneous interval is still open. Mercurio and Spokoiny (2004) have derived a homogeneity test for a supermartingale process. We show that a supermartingale of GH distributed variable can be obtained from the following lemma. It therefore leads to the same homogeneity test theory.

**LEMMA 1** *For every  $0 \leq \gamma \leq 1$  there exists a constant  $a_\gamma > 0$  such that*

$$\log \mathbb{E}[e^{u\zeta_\gamma}] \leq \frac{a_\gamma u^2}{2},$$

where  $\zeta_\gamma = (|\varepsilon|^\gamma - C_\gamma)/D_\gamma$  is the transformed GH distributed variable  $\varepsilon$ .

The proof of this lemma is given in the Appendix.

Consider a predictable process  $p_t$  (such as the volatility  $\sigma_t$  or the local parameter  $\theta_t$ ) w.r.t. the information set  $\mathcal{F}_{t-1}$ :

$$\Upsilon_t = \exp \left( \sum_{s=1}^t p_s \zeta_s - (a_\gamma/2) \sum_{s=1}^t p_s^2 \right)$$

$\Upsilon_t$  is a supermartingale, since

$$\begin{aligned} \mathbb{E}(\Upsilon_t | \mathcal{F}_{t-1}) - \Upsilon_{t-1} &= \mathbb{E}(\Upsilon_t | \mathcal{F}_{t-1}) - \mathbb{E}(\Upsilon_{t-1} | \mathcal{F}_{t-1}) \\ &= \mathbb{E} \left[ \exp \left( \sum_{s=1}^t p_s \zeta_s - (a_\gamma/2) \sum_{s=1}^t p_s^2 \right) \right. \\ &\quad \left. - \exp \left( \sum_{s=1}^{t-1} p_s \zeta_s - (a_\gamma/2) \sum_{s=1}^{t-1} p_s^2 \right) \middle| \mathcal{F}_{t-1} \right] \\ &= \mathbb{E} \left[ \exp \left( \sum_{s=1}^{t-1} p_s \zeta_s - (a_\gamma/2) \sum_{s=1}^{t-1} p_s^2 \right) \{ \exp(p_t \zeta_t - a_\gamma/2 p_t^2) - 1 \} \middle| \mathcal{F}_{t-1} \right] \\ &= \underbrace{\frac{\exp(p_1 \zeta_1)}{\exp(a_\gamma/2 p_1)}}_{\leq 1, \text{Lemma 1}} \cdots \underbrace{\frac{\exp(p_{t-1} \zeta_{t-1})}{\exp(a_\gamma/2 p_{t-1})}}_{\leq 1} \cdot \underbrace{\mathbb{E} \left[ \frac{\exp(p_t \zeta_t)}{\exp(a_\gamma/2 p_t)} - 1 \middle| \mathcal{F}_{t-1} \right]}_{\leq 1} \\ &\leq 0 \end{aligned}$$

i.e.  $\mathbb{E}(\Upsilon_t | \mathcal{F}_{t-1}) \leq \Upsilon_{t-1}$ . With this supermartingale property, the statistical properties of  $\hat{\theta}_I$  are given in:

**THEOREM 1** *If  $R_1, \dots, R_\tau$  obey the heteroscedastic model and the residual  $\varepsilon$  satisfies Lemma 1. Furthermore, the volatility coefficient  $\sigma_t$  satisfies the condition  $b \leq \sigma_t^2 \leq bB$  with some positive constants  $b$  and  $B$ , then it holds for the estimate  $\hat{\theta}_I$  of  $\theta_\tau$  :*

$$P\{|\hat{\theta}_I - \theta_\tau| > \Delta_I(1 + \eta s_\gamma |I|^{-1/2}) + \eta \hat{v}_I\} \leq 4\sqrt{e}\eta(1 + \log B) \exp\left\{-\frac{\eta^2}{2a_\gamma(1 + \eta s_\gamma |I|^{-1/2})^2}\right\}.$$

where  $\Delta_I$  is the squared bias defined as  $\Delta_I^2 = |I|^{-1} \sum_{t \in I} (\theta_t - \theta_\tau)^2$ .

Theorem 1 indicates that the estimation error  $|\hat{\theta}_I - \theta_\tau|$  is small relative to  $\eta \hat{v}_I$  for  $\tau \in I$  with a high probability, if  $I$  is a time homogeneous interval and therefore the squared bias  $\Delta_I$  is negligible. Straightforwardly, the following condition can be used to test the homogeneity hypothesis in an interval  $I$ :

$$|\hat{\theta}_I - \theta_\tau| \leq \eta \hat{v}_I.$$

In the test,  $I$  is split into two subintervals:  $I \setminus J$  and  $J$ . If  $I$  is a time homogeneous interval, the estimates based on the two subintervals must be very close. The homogeneity condition can be stated as:

$$|\hat{\theta}_{I \setminus J} - \hat{\theta}_J| \leq \eta(\hat{v}_{I \setminus J} + \hat{v}_J) = \eta'(\sqrt{\hat{\theta}_J^2 |J|^{-1}} + \sqrt{\hat{\theta}_{I \setminus J}^2 |I \setminus J|^{-1}}). \quad (14)$$

provided  $\eta' = \eta s_\gamma$  sufficiently large. If condition (14) is violated, the homogeneity hypothesis for the interval  $I$  is rejected.

The test procedure starts from an initial small interval  $I$  that satisfies the homogeneity and consists of 4 steps:

Step 1: Enlarge the interval  $I$  from  $[\tau - m_0, \tau)$  to  $[\tau - k \times m_0, \tau)$ , i.e.  $m = k \times m_0$ , and split the new interval into two subintervals  $J$  and  $I \setminus J$ . The parameters  $m_0$  and  $k$  are integers specified according to data. In this paper, we chose  $m_0 = 5$  and  $k = 2$ .

Step 2: Start homogeneity test for interval  $J = [\tau - \frac{2}{3}m, \tau)$ . If the homogeneity hypothesis isn't rejected, enlarge  $J$  one point further to  $[\tau - \frac{2}{3}m - 1, \tau)$  and repeat the homogeneity test (14). The loop will continue until the left point of the subinterval  $J$  reaches the point  $\tau - \frac{1}{3}m$ . The choice of  $\frac{1}{3}$  comes from the fact that the right  $\frac{1}{3}$  part has been tested in the last homogeneous interval and the left one-thirds will be tested in the next homogeneous interval, Mercurio and Spokoiny (2004).

Step 3: If (14) is violated at point  $s$ , the loop stops and the time homogeneous interval  $I$  is specified from point  $\tau$  to point  $s + 1$ .

Step 4: If time homogeneity holds for this interval, go back to Step 1.

The largest interval  $I$  is finally chosen as the time homogeneous interval for point  $\tau$ , based on

which the local volatility  $\sigma_\tau$  is estimated. However there are still two threshold parameters to be specified:  $\gamma$  in the power transformation and  $\eta'$  in the homogeneity test condition. According to Lemma 1, the parameter  $\gamma$  is bounded in  $[0, 1]$ . In our study, we chose  $\gamma = 0.5$  as same as the model based on the normal distribution to satisfy the comparability. The value of  $\eta'$  is similar to a smoothing parameter of the nonparametric regression. We thus propose a nonparametric way to pick up a global  $\eta'$ . Given a starting point  $t_0$  and provided that there are enough past observations to estimate  $\hat{\theta}_{(t_0, \eta')}$ , the value  $\eta'$  minimizes the forecast error:

$$\eta' = \operatorname{argmin} \sum_{t=t_0}^{\tau-1} \left\{ |R_t|^\gamma - \hat{\theta}_{(t, \eta')} \right\}^2. \quad (15)$$

### 2.3 Monte Carlo simulation

The GHADA technique consists of two main parts: estimate the GH distribution parameters, from which one calculates the quantile of the P&L, and predict the volatility using the adaptive methodology. The calculation procedure can be described as:

1. Select the transformation parameter  $\gamma$  and a starting point  $t_0$ .
2. Given different  $\eta'$ s, estimate the local volatilities using the adaptive methodology. Choose the  $\eta'$  which minimizes the forecast error and the corresponding estimated volatility process  $\hat{\sigma}_t$ .
3. Calculate the devolatilized returns  $\varepsilon_t = R_t / \hat{\sigma}_t$  and estimate the GH parameters.
4. Calculate the  $VaR_{t+1}$ , multiplying the volatility forecast  $\tilde{\sigma}_{t+1}$  and the quantile of the devolatilized return.

The previous calculations and the comparison in Section 2.1 have provided evidence that the GH distributions can represent the empirical distribution of the underlying well. In this section, we focus on the volatility estimation. In order to check the performance of the adaptive methodology, Monte Carlo simulation was applied. We intend to estimate the local volatility on the basis of simulated HYP and NIG variables and to analyze the sensitivity of the GHADA algorithm to jumps of the volatility. We considered two volatility processes:

$$\sigma_{1,t} = \begin{cases} 0.01 & : 1 \leq t \leq 400 \\ 0.05 & : 400 < t \leq 750 \\ 0.01 & : 750 < t \leq 1000 \end{cases} \quad (16)$$

$$\sigma_{2,t} = \begin{cases} |0.02t - 5| & : 1 \leq t \leq 300 \\ |0.02t - 10| & : 300 < t \leq 600 \\ |0.12t - 100| & : 600 < t \leq 1000 \end{cases} \quad (17)$$

We simulated 200 HYP and 200 NIG random variables. Each series consists of 1000 observations. We constructed the HYP return series by multiplying the HYP variables and the volatility  $\sigma_{1,t}$ . The NIG return series were constructed by multiplying the NIG trajectories and the volatility  $\sigma_{2,t}$ .

We applied the GHADA algorithm to estimate the local volatilities of these simulated returns. The first 200 observations were used to estimate local volatility at the starting point. The transformation parameter  $\gamma$  was pre-set as 0.5. The value  $\eta'$  that minimized the forecast error was selected and used in the homogeneity test. We repeated the multiple homogeneity test and specified the homogeneity intervals for the points from 201 to 1000. Two examples of the estimated local constant volatility series - one belongs to the HYP distributed returns and the other the NIG returns - are displayed in Figure 4. The estimated volatility processes for the simulated 200 HYP and 200 NIG distributed returns can be downloaded at <http://ise.wiwi.hu-berlin.de/~ychen/ghada/> : simulation1.AVI and simulation2.AVI respectively. Compared to the true volatility processes (dashed line), these estimates (straight line) satisfactorily represent the movements and the sudden changes of the volatility process. To study the sensitivity of our approach, we introduced a percentage rule that told us after how many steps the sudden volatility jump is detected at 40%, 50% or 60% level of the jump size. The 40% rule, for example, refers to the number of time steps necessary to reach 40% of the jump size. Table 1 gives an overview of the detection delays. One interesting feature of the GHADA approach is its out-performance of catching up a sudden increasing jump in the volatility process. Concerning the volatility process  $\sigma_1$ , the first jump (occurring at the 401<sup>st</sup> point) was at 40% level on average detected after 6 steps. This number increases to 8 for the 60% rule of the jump. In contrast, the detection speed is slow for downward jumps. For the second jump at  $t = 750$ , where the volatility decreases from 5% to 1%, the yields 40% rule about 12 steps, twice slower reacting than that to the increasing jump with the same quantity. This fact results from a loose test power. In the homogeneity test (14), the squared conditional variance  $v_I$  depends on  $\theta_t$ , a larger value of  $\theta_t$  thus induces a loose test power. Consequently, a slower detection speed occurs in the jump down from a high value. Additionally, we considered two jumps in the second volatility process  $\sigma_2$ . The mean values of the steps detecting 40%, 50% and 60% of these two jumps at  $t = 300$  and  $t = 600$  are even smaller than those in  $\sigma_1$ .

The mean squared error (MSE) of the estimation  $\frac{1}{800} \sum_{t=201}^{1000} (\hat{\sigma}_{1,t} - \sigma_{1,t})^2$  is  $3.26(10^{-5})$  for the simulation 115 (HYP). Figure 5 illustrates the mean process and the 99% confidence interval of the estimates. At the jump point the estimate is turbulent and the confidence interval increases. In the NIG case, the volatility process ( $\sigma_2$ ) is more volatile and not be constant in a short interval any more. Nevertheless the GHADA model gave fine results and the big changes of the volatility movement were caught in about 6 points as well. The MSE of the simulation example is 6.88. From Figure 6 one can further see that the mean process repeats the movement of the volatility process. The confidence interval around the



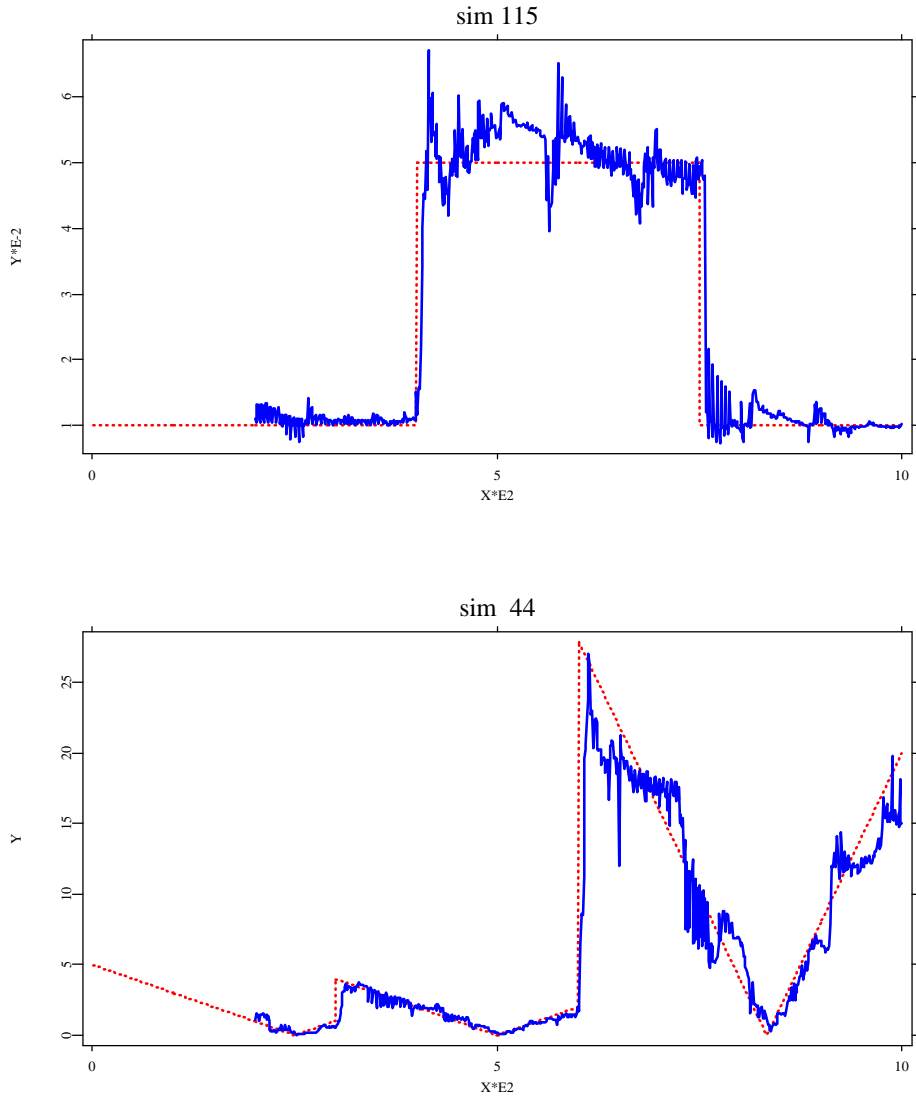


Figure 4: The estimated volatility processes on the basis of two simulated examples (dashed line): sim 115 (HYP) and sim 44 (NIG). The power transformation parameter  $\gamma = 0.5$  and the starting point  $t_0 = 201$ .

[GHADAsim1.xpl](#) [GHADAsim2.xpl](#)

biggest change is larger than the HYP example. The scale of the volatility process could be one reason for the large span.

	mean	standard deviation	maximum	minimum
Detection decay to the first jump at $t = 400 - \sigma_1$				
40% rule	5.9	2.4	15	1
50% rule	6.9	2.6	19	2
60% rule	7.9	2.9	19	2
Detection decay to the second jump at $t = 750 - \sigma_1$				
40% rule	11.8	4.4	39	3
50% rule	13.5	6.5	58	5
60% rule	15.9	10.9	98	6
Detection decay to the first jump at $t = 300 - \sigma_2$				
40% rule	4.9	2.4	13	0
50% rule	6.2	3.0	18	2
60% rule	7.6	4.2	33	2
Detection decay to the first jump at $t = 600 - \sigma_2$				
40% rule	4.7	1.9	12	0
50% rule	5.7	2.7	23	2
60% rule	6.8	3.4	24	2

Table 1: Descriptive statistics for the detecting speeds to the sudden jumps of the volatility processes.

### 3 RISK MANAGEMENT

#### 3.1 Data Set

Two data sets: DEM/USD exchange rate and a German bank portfolio were used in our empirical analysis. They are available at MD\*Base ([www.quantlet.org/mdbase](http://www.quantlet.org/mdbase)).

The exchange rate is daily registered from 1979/12/01 to 1994/04/01. There are 3720 observations. The first 500 observations are used as a basis to estimate the local volatility and the GH distribution parameters. The bank portfolio data reports the market value of the portfolio holden by a German bank. There are 5603 observations. Figure 1 and Figure 3 have shown the empirical (log) densities and the estimated HYP and NIG densities of the exchange rate data. We estimated the GH distribution parameters using the maximum likelihood method. Using the proposed GHADA approach, the local volatility estimates of the DEM/USD exchange rates are displayed in Figure 7 where  $t_0$  is 501 and  $\eta' = 1.06$ . Figure 8 and Figure 9 show the respective estimates of the bank portfolio data.

The devolatilized return

$$\hat{\varepsilon}_t = R_t / \hat{\sigma}_t \quad (18)$$

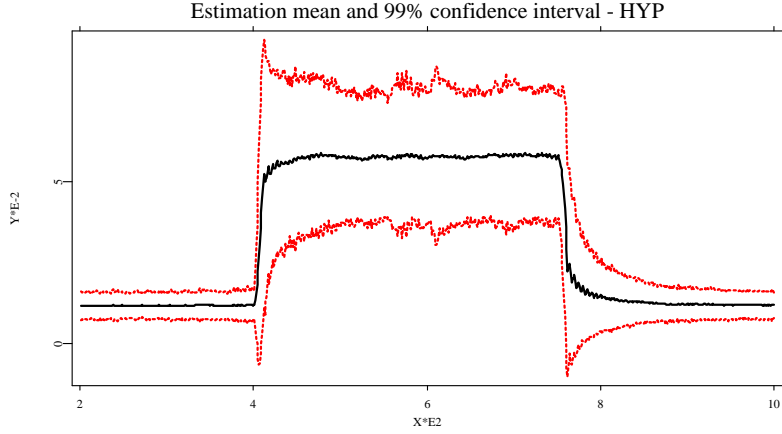


Figure 5: The mean process of the local volatility estimates based on 200 simulations and the 99% point confidence interval (dashed line) of the estimation.

[GHADAsim1.xpl](#)

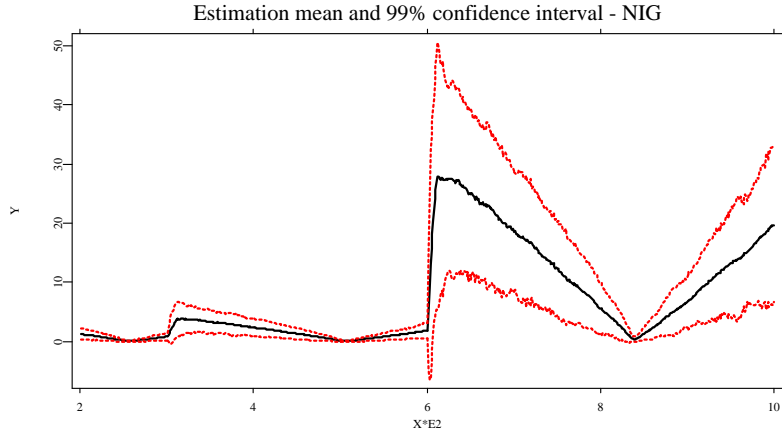


Figure 6: The mean process of the local volatility estimates based on 200 simulations and the 99% point confidence interval (dashed line) of the estimation.

[GHADAsim2.xpl](#)

is expected to be i.i.d. but not necessarily normally distributed. We assume that the devolatilized return is HYP or NIG distributed. Table 2 summarizes the descriptive statistics of the devolatilized returns for the exchange rate data and the bank portfolio data. The first and the second moments are close to 0 and 1, but the kurtoses of these two data sets are high with respect to the normal distribution. The graphics of the devolatilized return processes are displayed in Figure 7 and Figure 9. Figure 10 shows the boxplots of the time homogeneous intervals' length for these two data sets. The average interval length for the exchange rate data is 51 while the length for the German bank portfolio data is 71. It can be seen that the spans of the interval length are wide for two data sets. It provides an evidence

	exchange rate data	bank portfolio data
mean	-0.0052	0.0113
std	0.9938	0.9264
skewness	-0.0121	-0.0815
kurtosis	4.0329	5.1873

Table 2: Descriptive statistics for the devolatilized residuals of the exchange rate data and bank portfolio data.

that the volatility model changes from time to time. Compared to a fixed time-invariant model, the adaptive model is flexible and simple to catch the fluctuation of the volatility.

### 3.2 Value at Risk

Value at risk (VaR) is one of the most often used risk measure. It measures the possible loss level over a given horizon at a given confidence level  $1 - p$  and answers the question: How much can I lose with  $p$  probability over the pre-set horizon. The research on VaR models has been ignited and prompted by the rule of Basel Committee on Banking Supervision in 1995: financial institutions may use their internal VaR models. The selection of the internal VaR model as well as the volatility estimation is essential to the VaR based risk management. Let  $q_p$  denote the  $p$ -th quantile of the distribution of  $\varepsilon_t$ , i.e.  $P(\varepsilon_t < q_p) = p$ , we have  $P(R_t < \sigma_t q_p | \mathcal{F}_{t-1}) = p$ . Mathematically the VaR is defined as:

$$\text{VaR}_{p,t} = F_t^{-1}(p) = \sigma_t q_p.$$

where  $F_t^{-1}$  is the inverse function of the conditional cumulative distribution function of the underlying at time  $t$ , Franke et al. (2004).

In practice, we are interested in the forecast of VaR. Using the GHADA approach, we adaptively estimated the volatility  $\hat{\sigma}_t$ . Since the volatility process is a supermartingale, it is natural to use the estimate today as the volatility forecast  $\tilde{\sigma}_{t+1}$  for tomorrow, i.e.  $\tilde{\sigma}_{t+1} = \hat{\sigma}_t$ . Furthermore we estimated the HYP and NIG distribution parameters of the devolatilized returns and calculated the quantile  $q_p$ . The VaR at the probability level  $p$  was forecasted as:

$$\tilde{\text{VaR}}_{p,t+1} = \tilde{\sigma}_{t+1} \hat{q}_p.$$

As same as the volatility model, the distribution parameters could be time-variant as well. Figure 11 shows the HYP-quantile forecasts based on the 500 past devolatilized returns of the exchange rate for each time point. It provides an evidence that the quantile varies as time passes, especially for extreme probability levels such as  $p = 0.005$ . In this context we

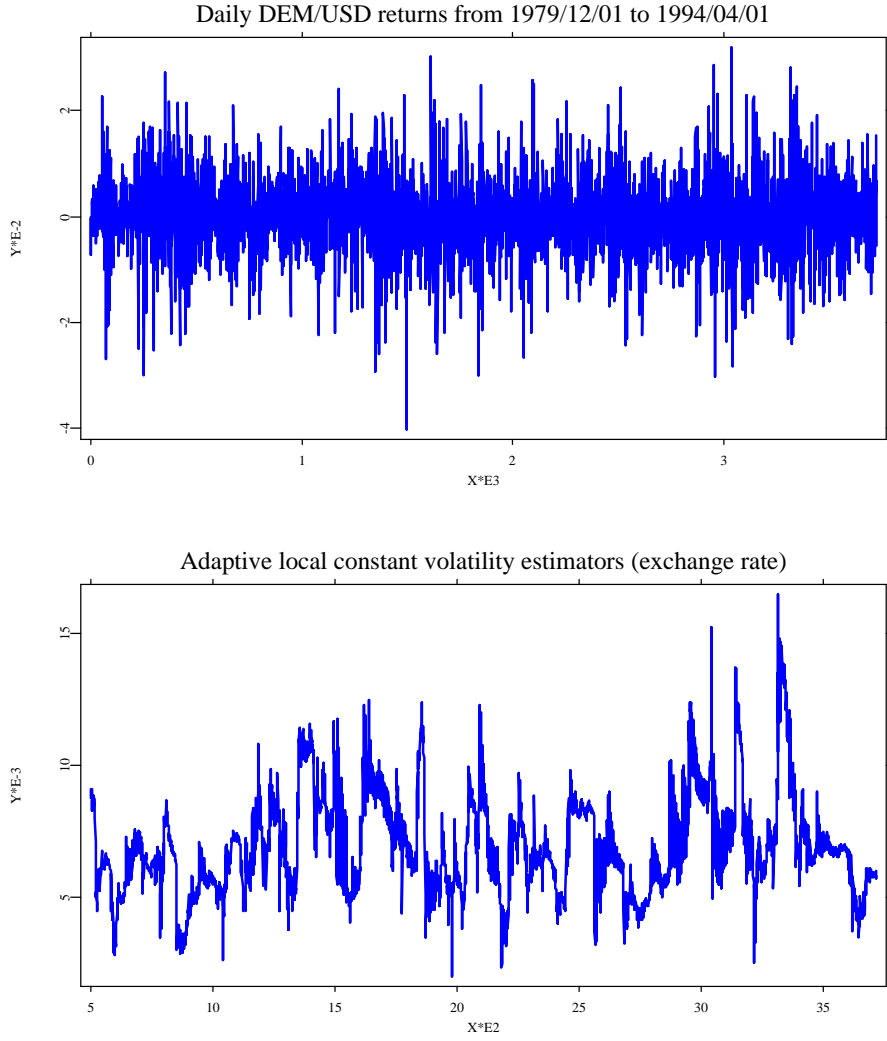


Figure 7: The return process of DEM/USD exchange rates (upper) and its adaptive volatility estimates (lower) for  $t_0 = 501$  and  $\eta' = 1.06$ .

 GHADAfx.xpl

couldn't stick to the assumption that the innovations are identically distributed. Instead, we updated the distribution parameters daily based on the previous 500 data points. The same phenomenon holds for NIG distribution, which is omitted here.

The observations whose losses exceed the VaR are called exceptions. The daily VaR forecasts of the DEM/USD rates are displayed in Figure 12 and Figure 13. The VaR forecasts are different between the GHADA model and the model with the normal distribution (normal model). As we have discussed, the VaRs based on the normal distribution are almost identical to the HYP distribution at the 5% probability level. Sometime the normal model performs even better than the GHADA model. However at the 5% level, there are more than 169 exceptions observed in 17 years, i.e. more than 12 exceptions annually. In order to control the risk exposure to the market, the extreme events (lower levels) should

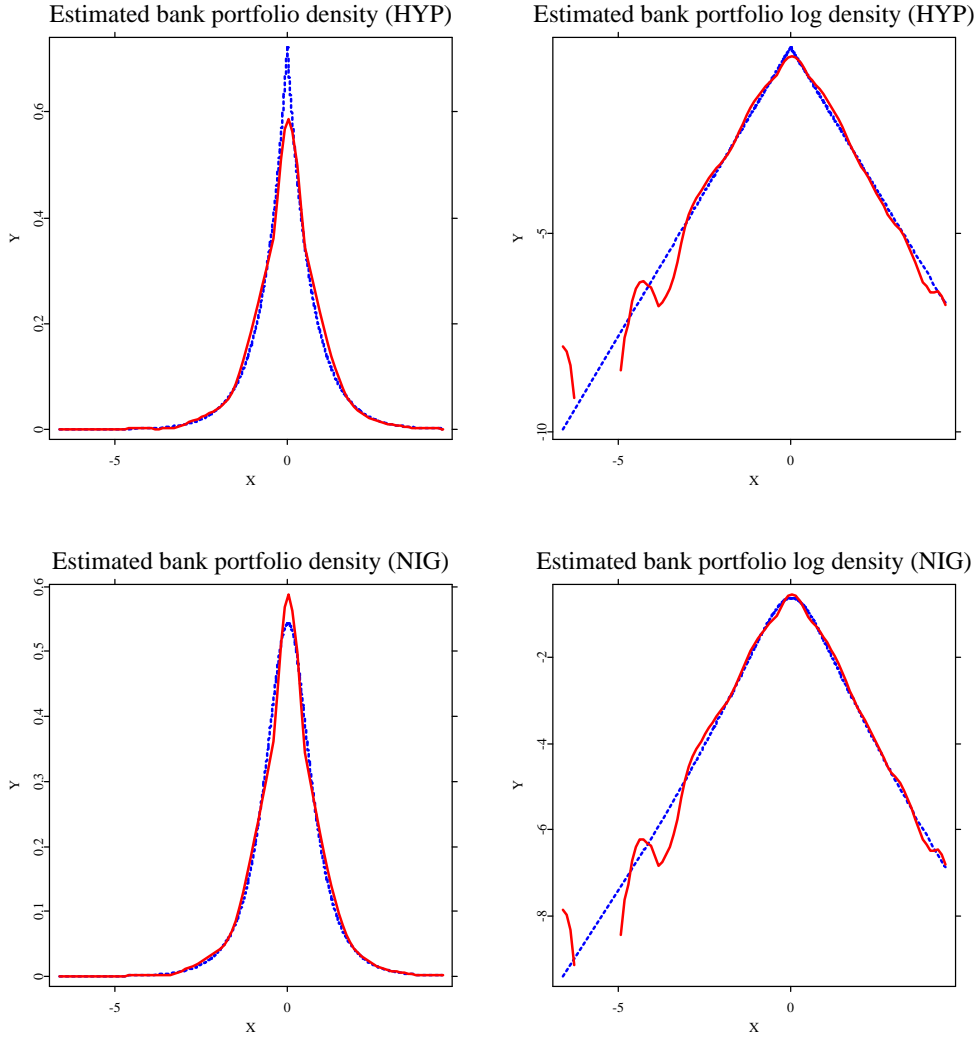


Figure 8: The estimated density (left) and log density (right) of the standardize return of the German bank portfolio rates (red) with nonparametric kernel ( $h \approx 0.61$ ) and the estimated HYP density (blue) on the top with the maximum likelihood estimators  $\hat{\alpha} = 1.819$ ,  $\hat{\beta} = -0.168$ ,  $\hat{\delta} = 0.705$ ,  $\hat{\mu} = 0.145$  and the estimated NIG density (blue) with the maximum likelihood estimators  $\hat{\alpha} = 1.415$ ,  $\hat{\beta} = -0.171$ ,  $\hat{\delta} = 1.254$  and  $\hat{\mu} = 0.146$ .

 [GHADAkupfer.xpl](#)

also be considered. It is obvious that as the probability level decreases to some extreme values such as 1% or 0.5%, the gaps of these two models get larger and larger. Except in the 5% case, the GHADA model is superior to the normal model at the other three levels of 2.5%, 1% and 0.5%.

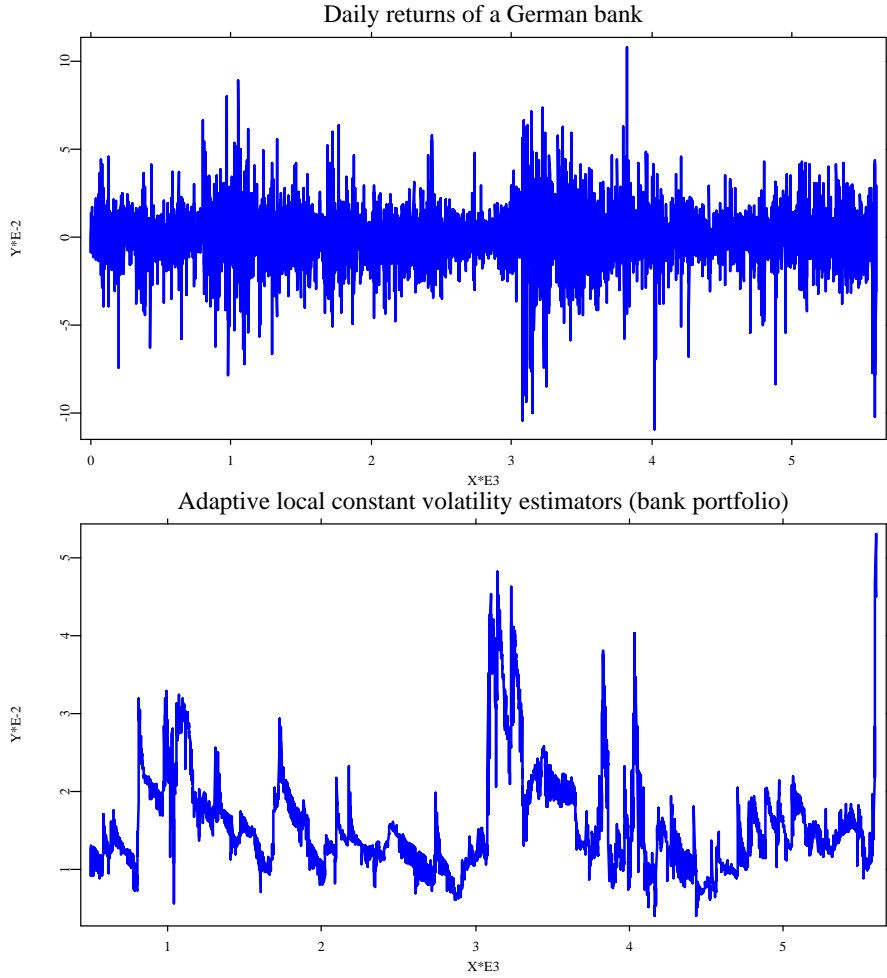


Figure 9: The return process of a German bank's portfolio (upper) and its adaptive volatility estimates (lower) for  $t_0 = 501$  and  $\eta' = 1.23$ . The average length of time homogeneous interval is 71.

 [GHADAkupfer.xpl](#)

### 3.3 Backtesting VaR

To evaluate the validation of the VaR calculation, consider the backtesting procedures presented in Christoffersen (1998). Above all, a VaR calculation should not underestimate the market risk. Let  $1_t$  denote the indicator of exceptions at time point  $t$ ,  $t = 1, 2, \dots, T$ . If the proportion of exceptions,  $N/T = T^{-1} \sum_{t=1}^T 1_t$ , is much larger than  $p$ , i.e.  $N/T > p$ , it means that the possible losses happen more often than the fixed level. In this case, the VaR model should be rejected. One constructs a hypothesis to test:

$$H_0 : \mathbb{E}[N] = Tp \quad \text{vs.} \quad H_1 : \mathbb{E}[N] \neq Tp \quad (19)$$

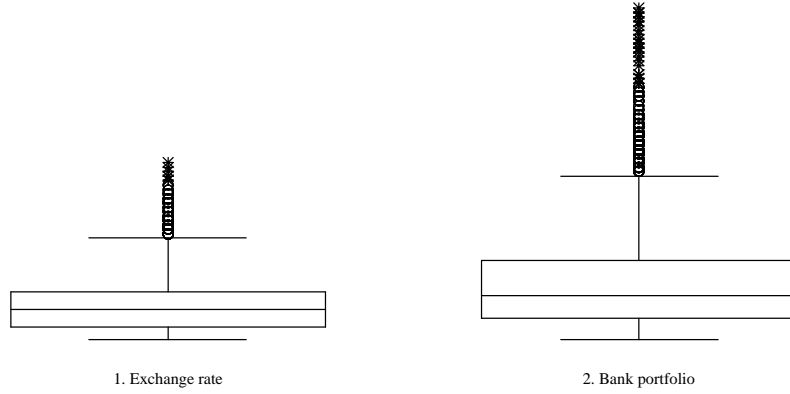


Figure 10: Boxplots of the DEM/USD exchange rates (left) and the German bank portfolio data (right).

Under  $H_0$ ,  $N$  is a Binomial random variable with parameters  $T$  and  $p$ , the likelihood ratio test statistic can be derived as:

$$\text{LR1} = -2 \log \{(1-p)^{T-N} p^N\} + 2 \log \{(1-N/T)^{T-N} (N/T)^N\}, \quad (20)$$

which is asymptotically  $\chi^2(1)$  distributed, Jorion (2001).

In addition, a VaR model that yields exception clusters should also be rejected. Since a cluster of VaR exceedances means that if there is an exception today, an exception may also occur tomorrow with a higher probability than the prescribed level  $p$ . Another important test is the test of independence. Let us denote  $\pi_{ij} = P(1_t = j | 1_{t-1} = i)$  as the transition probability and  $n_{ij} = \sum_{t=1}^T 1(1_t = j \text{ and } 1_{t-1} = i)$ , where  $i, j = 0$  or  $1$ . The independence hypothesis is given as:

$$H_0 : \pi_{00} = \pi_{10} = \pi, \pi_{01} = \pi_{11} = 1 - \pi \quad (21)$$

One can test this hypothesis using the likelihood ratio statistic:

$$\text{LR2} = -2 \log \{\hat{\pi}^{n_0} (1 - \hat{\pi})^{n_1}\} + 2 \log \{\hat{\pi}_{00}^{n_{00}} \hat{\pi}_{01}^{n_{01}} \hat{\pi}_{10}^{n_{10}} \hat{\pi}_{11}^{n_{11}}\}, \quad (22)$$

where  $\hat{\pi}_{ij} = n_{ij} / (n_{ij} + n_{i,1-j})$ ,  $n_j = n_{0j} + n_{1j}$ , and  $\hat{\pi} = n_0 / (n_0 + n_1)$ . Under  $H_0$ , LR2 is asymptotically  $\chi^2(1)$  distributed as well, Jorion (2001).

Table 3 and Table 4 summarize the results of the backtesting for the DEM/USD data and the bank portfolio data. On average, the GHADA model gives more accurate forecasts at each probability level than the normal model. For example, the proportions of exceptions



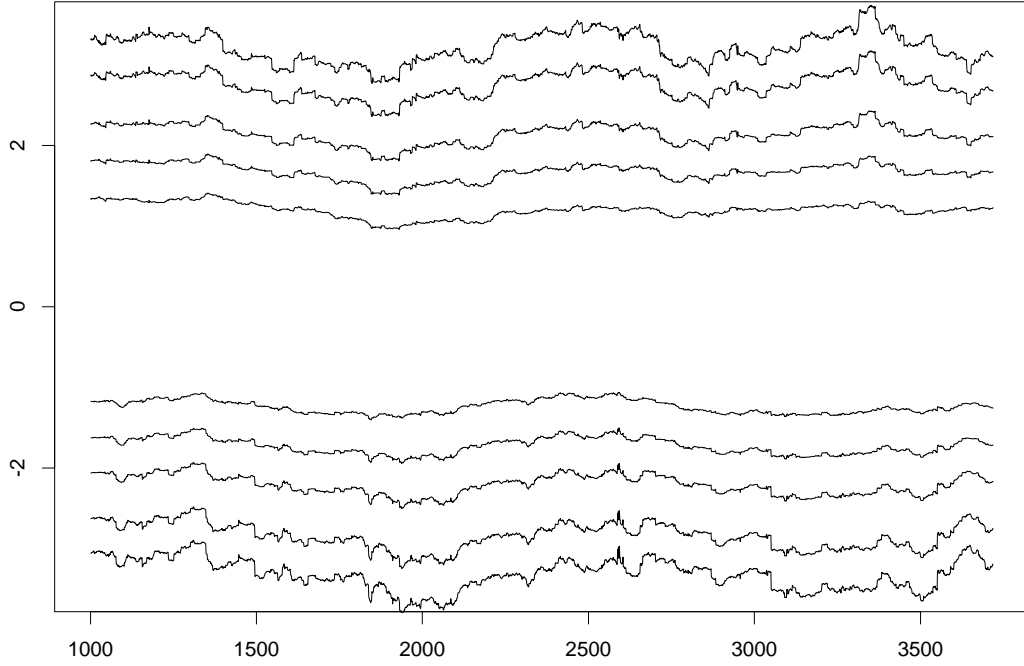


Figure 11: Quantiles estimated based on the past 500 devolatilized returns of the exchange rate. From the top the evolving HYP quantiles for  $p = 0.995$ ,  $p = 0.99$ ,  $p = 0.975$ ,  $p = 0.95$ ,  $p = 0.90$ ,  $p = 0.10$ ,  $p = 0.05$ ,  $p = 0.025$ ,  $p = 0.01$ ,  $p = 0.005$ .

at 1% level relative to the normal model is about 1.5%, which is one and a half times of these same level indices of GHADA models. The test of VaR level is not rejected for HYP and NIG based model at all levels. In contrast, the normal model fails to provide acceptable results at the extreme levels. In addition, all these models fulfill the independence test.

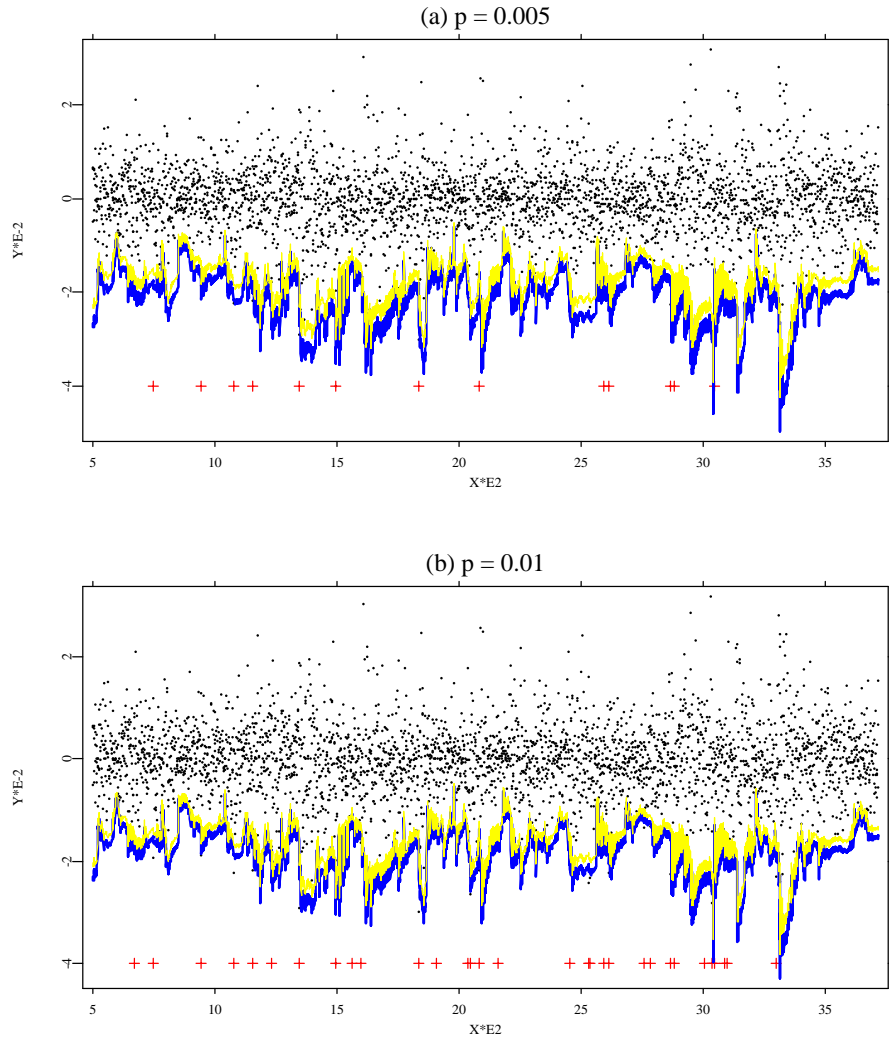



Figure 12: Value at Risk forecast plots for DEM/USD data. The dots are the returns, the solid line is the VaR forecast based on HYP underlying distribution, the yellow line is the VaR forecast with normal distribution, and the crosses indicate the VaR exceptions of HYP model. (a)  $p = 0.005$ . (b)  $p = 0.01$ .

 GHADAFXVAR.xpl

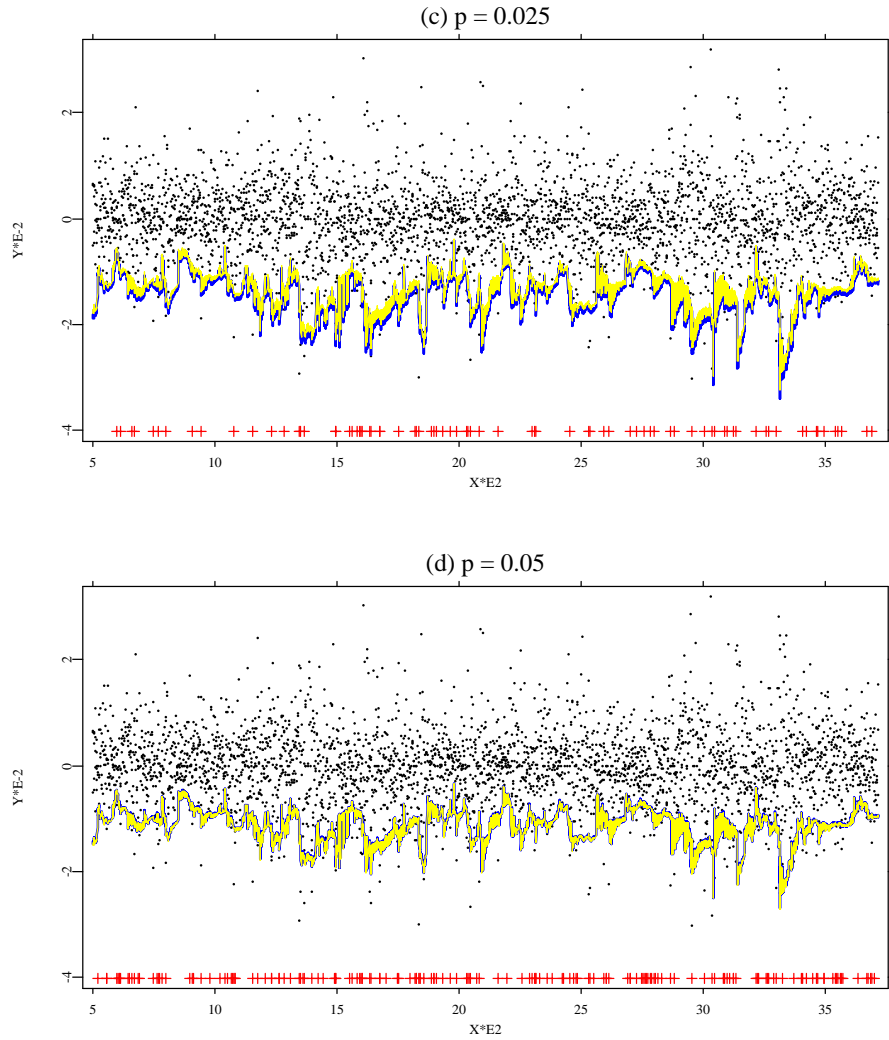



Figure 13: Value at Risk forecast plots for DEM/USD data. The dots are the returns, the solid line is the VaR forecast based on HYP underlying distribution, the yellow line is the VaR forecast with normal distribution, and the crosses indicate the VaR exceptions of HYP model. (c)  $p = 0.025$ . (d)  $p = 0.05$ .

[GHADafxvar.xpl](#)

Model	$p$	$N/T$	LR1	p-value	LR2	p-value
Normal	0.005	0.01025	13.667	0.000*	0.735	0.391
	0.01	0.01460	6.027	0.014	0.138	0.710
	0.025	0.02858	1.619	0.203	0.056	0.813
	0.05	0.05250	0.417	0.518	0.007	0.934
HYP	0.005	0.00403	0.640	0.424	0.189	0.664
	0.01	0.00963	0.045	0.832	0.655	0.419
	0.025	0.02485	0.003	0.957	0.666	0.415
	0.05	0.05312	0.648	0.421	0.008	0.927
NIG	0.005	0.00404	0.640	0.424	0.189	0.664
	0.01	0.00994	0.001	0.973	0.694	0.405
	0.025	0.02516	0.004	0.953	0.719	0.396
	0.05	0.05405	1.086	0.297	0.040	0.841

Table 3: Backtesting results for DEM/USD example. \* indicates the rejection of the model which is used.

 [GHADafxvar.xpl](#)

Model	$p$	$N/T$	LR1	p-value	LR2	p-value
Normal	0.005	0.010	19.809	0.000*	1.070	0.301
	0.01	0.016	13.278	0.000*	0.422	0.516
	0.025	0.028	2.347	0.126	0.781	0.377
HYP	0.005	0.003	5.111	0.024	0.160	0.689
	0.01	0.008	2.131	0.144	0.705	0.401
	0.025	0.025	0.053	0.819	1.065	0.302
NIG	0.005	0.003	5.111	0.024	0.160	0.689
	0.01	0.009	0.747	0.387	0.841	0.359
	0.025	0.027	0.438	0.508	1.429	0.232

Table 4: Backtesting results for the bank portfolio example. \* indicates the rejection of the model which is used.

## 4 FINAL REMARKS

We have proposed a risk management (GHADA) model based on the adaptive volatility estimation and the generalized hyperbolic distribution. Our study is summarized as follows.

- The adaptive volatility estimation methodology by Mercurio and Spokoiny (2004) is also applicable with generalized hyperbolic distribution. The threshold parameter used to specify the time homogeneity interval can be estimated in a nonparametric way.
- The distribution of the devolatilized returns from the adaptive volatility estimation is found to be leptokurtic and, sometimes, asymmetric. We found that the distribution of the innovations can be perfectly modelled by the HYP and NIG distributions, subclasses of the generalized hyperbolic distribution.
- The proposed approach can be easily applied to calculate and forecast risk measures such as value at risk and expected shortfall. On the basis of the DEM/USD data and a German bank portfolio data it shows that the proposed approach performs better than a model with the normal distribution.

In financial markets, it is more interesting and challenging to measure the risk levels of multiple time series. Härdle, Herwartz and Spokoiny (2003) proposed an idea of the adaptive volatility estimation method for a multiple time series. We expect to apply the idea to a model with the multivariate generalized hyperbolic distribution. A detailed study on this approach is left for our future research.

## 5 APPENDIX

**Proof of Lemma 1.**

**Proof:**

Firstly we show that the moment generating function  $\mathbb{E}[e^{u\zeta_\gamma}]$  exists for all  $u \in R$ .

Suppose that  $\mathcal{L}(x) = GH(\lambda, \alpha, \beta, \delta, \mu)$  with the density function  $f$  for the transformed variable  $y \stackrel{\text{def}}{=} |x|^\gamma$ , we have

$$P(y \leq z) = P(-z^{\frac{1}{\gamma}} \leq x \leq z^{\frac{1}{\gamma}}) = \int_{-\infty}^{z^{\frac{1}{\gamma}}} f(x)dx - \int_{-\infty}^{-z^{\frac{1}{\gamma}}} f(x)dx, \quad z > 0$$

Then the density of  $y \in (0, \infty)$  is:

$$\begin{aligned} g(z) = \frac{d}{dz} P(y \leq z) &= \gamma^{-1} \{f(z^{\frac{1}{\gamma}})z^{\frac{1}{\gamma}-1} + f(-z^{\frac{1}{\gamma}})z^{\frac{1}{\gamma}-1}\} \\ &= \gamma^{-1} z^{\frac{1}{\gamma}-1} \{f(z^{\frac{1}{\gamma}}) + f(-z^{\frac{1}{\gamma}})\}, \quad z > 0. \end{aligned}$$

Since  $f_{GH}(x; \lambda, \alpha, \beta, \delta, \mu = 0) \sim x^{\lambda-1} e^{-(\alpha-\beta)x}$  as  $x \rightarrow \pm\infty$ , it follows

$$\begin{aligned} g(z) &\sim \frac{z^{\frac{1}{\gamma}-1}}{\gamma} \{z^{\frac{\lambda-1}{\gamma}} e^{(\beta-\alpha)z^{\frac{1}{\gamma}}} + z^{\frac{\lambda-1}{\gamma}} e^{-(\beta+\alpha)z^{\frac{1}{\gamma}}}\} \\ &= \frac{z^{\frac{\lambda}{\gamma}-1}}{\gamma} \{e^{(\beta-\alpha)z^{\frac{1}{\gamma}}} + e^{-(\beta+\alpha)z^{\frac{1}{\gamma}}}\}, \quad z \rightarrow \infty \end{aligned}$$

For  $\gamma < 1$ , it holds that  $\int_0^\infty e^{uz} g(z) dz < \infty \quad \forall u \in R$ , since

$$\begin{aligned} \lim_{z \rightarrow \infty} (\beta - \alpha)z^{\frac{1}{\gamma}} + uz &\rightarrow -\infty \quad \forall u \in R \\ \lim_{z \rightarrow \infty} -(\beta + \alpha)z^{\frac{1}{\gamma}} + uz &\rightarrow -\infty \quad \forall u \in R \end{aligned}$$

Since the integration depends only on the exponential part, it holds also that

$$\int_0^\infty z^n e^{uz} g(z) dz = \int_0^\infty \frac{\partial^n}{\partial u^n} (e^{uz}) g(z) dz = \frac{\partial^n}{\partial u^n} \mathbb{E}[e^{uy}] < \infty,$$

then it can be shown that the moment generating function and  $\log(\mathbb{E}[e^{uy}])$  are smooth. It holds for every  $t > 0$ ,

$$\begin{aligned} \mathbb{E}[e^{uy}] = \mathbb{E}[e^{u|x|^\gamma}] &= \mathbb{E}[e^{u|x|^\gamma} 1(|x| \leq t)] + \mathbb{E}[e^{u|x|^\gamma} 1(|x| > t)] \\ &\leq e^{ut^\gamma} + \mathbb{E}[e^{|x|ut^{\gamma-1}} I(|x| > t)], \end{aligned} \tag{23}$$

Without loss of generality, we assume  $\mu = 0$ . Further

$$f_{GH}(x; \lambda, \alpha, \beta, \delta, \mu = 0) \sim x^{\lambda-1} e^{-(\alpha-\beta)x} \text{ as } x \rightarrow \infty,$$

and  $\int_y^\infty x^{\lambda-1} e^{-x} dx \sim y^{\lambda-1} e^{-y}$  as  $y \rightarrow \infty$ , Press, Teukolsky, Vetterling and Flannery (1992).

For an arbitrary but fixed  $u \in R_+$  and  $t_0 > 1$  so that  $ut^{\gamma-1} < \alpha - \beta$ , it holds for all  $t \geq t_0$

$$\begin{aligned} f(t) &\leq C_1 t^{\lambda-1} e^{(\beta-\alpha)t} \\ \int_{(\alpha-\beta-ut^{\gamma-1})t}^\infty x^{\lambda-1} e^{-x} dx &\leq C_2 [(\alpha - \beta - ut^{\gamma-1})t]^{\lambda-1} e^{-(\alpha-\beta-ut^{\gamma-1})t} \end{aligned}$$

where  $C_1, C_2 > 1$ .

Consequently for  $t \geq t_0$ ,

$$\begin{aligned}
\mathbb{E}[e^{u|t|^{\gamma-1}x}1(|x| > t)] &= \int_t^\infty e^{ut^{\gamma-1}x}f(x)dx \leq C_1 \int_t^\infty e^{ut^{\gamma-1}x}x^{\lambda-1}e^{-(\alpha-\beta)x}dx \\
&= C_1 \int_t^\infty x^{\lambda-1}e^{-(\alpha-\beta-ut^{\gamma-1})x}dx \\
&= C_1(\alpha-\beta-ut^{\gamma-1})^{-\lambda} \int_{(\alpha-\beta-ut^{\gamma-1})t}^\infty x^{\lambda-1}e^{-x}dx \\
&\leq C_1C_2t^{\lambda-1}e^{-(\alpha-\beta-ut^{\gamma-1})t}(\alpha-\beta-ut^{\gamma-1}t)^{-1}
\end{aligned} \tag{24}$$

If  $u$  is so large that  $t \stackrel{\text{def}}{=} (\frac{\alpha-\beta}{2})^{\frac{1}{\gamma-1}}u^c \geq t_0$  with  $\frac{1}{1-\gamma} \leq c$ , then (24) holds true since  $ut^{\gamma-1} = (\frac{\alpha-\beta}{2})uu^{c(\gamma-1)} \leq \frac{\alpha-\beta}{2} < \alpha-\beta$ .

Given  $t = (\frac{\alpha-\beta}{2}u)^{\frac{1}{1-\gamma}}$ , we get

$$\mathbb{E}[e^{ut^{\gamma-1}x}1(|x| > t)] \leq \frac{2C_1C_2}{\alpha-\beta}(\frac{\alpha-\beta}{2}u)^{\frac{\lambda-1}{1-\gamma}}e^{-\frac{\alpha-\beta}{2}(\frac{\alpha-\beta}{2}u)^{\frac{1}{1-\gamma}}}.$$

From which we get

$$\log(\mathbb{E}[e^{ut^{\gamma-1}x}1(x > t)]) \leq C_3 + \frac{\lambda-1}{1-\gamma} \log(u) - (\frac{\alpha-\beta}{2})^{\frac{2-\gamma}{1-\gamma}}u^{\frac{1}{1-\gamma}}$$

Further  $\log(\mathbb{E}[e^{ut^{\gamma-1}x}1(x > t)])u^{-\frac{1}{1-\gamma}}$  is also bounded for  $u \rightarrow \infty$ . Analogously we can show the bounding of  $\log(\mathbb{E}[e^{ut^{\gamma-1}x}1(x < -t)])u^{-\frac{1}{1-\gamma}}$ . Therefore for  $\gamma < 1$  the whole term  $\mathbb{E}[e^{u|x|^\gamma}1(|x| > t)]u^{-\frac{1}{1-\gamma}}$  is bounded as  $u \rightarrow \infty$ .

Given  $t = (\frac{\alpha-\beta}{2}u)^{\frac{1}{1-\gamma}}$ , we have

$$\begin{aligned}
e^{ut^\gamma} &= e^{(\frac{\alpha-\beta}{2})^{\frac{\gamma}{1-\gamma}}u^{\frac{1}{1-\gamma}}} \\
u^{-\frac{1}{1-\gamma}} \log(e^{ut^\gamma}) &= (\frac{\alpha-\beta}{2})^{\frac{\gamma}{1-\gamma}} = \text{constant}
\end{aligned}$$

Thus  $u^{-\frac{1}{1-\gamma}} \log(\mathbb{E}[e^{u|x|^\gamma}]) \leq u^{-\frac{1}{1-\gamma}} [\log(e^{ut^\gamma}) + \log\{E[e^{ut^{\gamma-1}|x|}1(|x| > t)]\}]$  is bounded for  $u \rightarrow \infty$ , i.e. for a sufficient large  $u_0$  there exist a constant  $C_u > 0$  such that

$$\mathbb{E}[e^{u|x|^\gamma}] \leq C_u u^{\frac{1}{1-\gamma}}, \quad u \geq u_0.$$

□

## References

- Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size, *Proceedings of the Royal Society of London A* **353**: 401–419.
- Barndorff-Nielsen, O. (1997). Normal inverse gaussian distributions and stochastic volatility modelling, *Scandinavian Journal of Statistics* **24**: 1–13.
- Barndorff-Nielsen, O. E. and Blæsild (1981). *Hyperbolic distribution and ramifications: Contributions to theory and Applications*, Vol. 4 of *Statistical Distributions in Scientific Work*, D. Reidel, pp. 19–44.
- Bibby, B. M. and Sørensen, M. (2001). *Hyperbolic Processes in Finance*, Technical Report 88, University of Aarhus, Aarhus School of Business.
- Bollerslev, T. (1995). *Generalised autoregressive conditional heteroskedasticity*, ARCH, selected readings, Oxford University Press, pp. 42–60.
- Christoffersen, P. F. (1998). Evaluating interval forecast, *International Economic Review* **39**: 841–862.
- Eberlein, E. and Keller, U. (1995). Hyperbolic distributions in finance, *Bernoulli* **1**: 281–299.
- Eberlein, E., Kallsen, J. and Kristen, J. (2003). Risk management based on stochastic volatility, *Journal of Risk* **5**: 19–44.
- Engle, R. F. (1995). *Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation*, ARCH, Oxford University Press.
- Franke, J., Härdle, W. and Hafner, C. (2004). *Statistics of Financial Markets*, Springer-Verlag Berlin Heidelberg New York.
- Härdle, W., Herwartz, H. and Spokoiny, V. (2003). Time inhomogeneous multiple volatility modelling, *Journal of Financial Econometrics* **1**: 55–95.
- Härdle, W., Müller, M., Sperlich, S. and Werwatz, A. (2004). *Nonparametric and Semiparametric Models*, Springer-Verlag Berlin Heidelberg New York.
- Harvey, A., R. E. and Shephard, N. (1995). *Multivariate stochastic variance models*, ARCH, selected readings, Oxford University Press, pp. 256–276.
- Jaschke, S. and Jiang, Y. (2002). Approximating value at risk in conditional gaussian models, in W. Härdle, T. Kleinow and G. Stahl (eds), *Applied Quantitative Finance*, Springer-Verlag Berlin Heidelberg New York.
- Jorion, P. (2001). *Value at Risk*, McGraw-Hill.



- Mercurio, D. and Spokoiny, V. (2004). Statistical inference for time inhomogeneous volatility models, *Annals of Statistics* **32**: 577–602.
- Press, W., Teukolsky, S., Vetterling, W. and Flannery, B. (1992). *Numerical Recipes in C*, Cambridge University Press.

## SFB 649 Discussion Paper Series

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Nonparametric Risk Management with Generalized Hyperbolic Distributions" by Ying Chen, Wolfgang Härdle and Seok-Oh Jeong, January 2005.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

