Nonparametric Risk Management with Generalized Hyperbolic Distributions

Ying Chen
Wolfgang Härdle

Center for Applied Statistics and Economics
Institut für Statistik and Ökonometrie
Humboldt-Universität zu Berlin

A Stylized Fact in Financial Markets

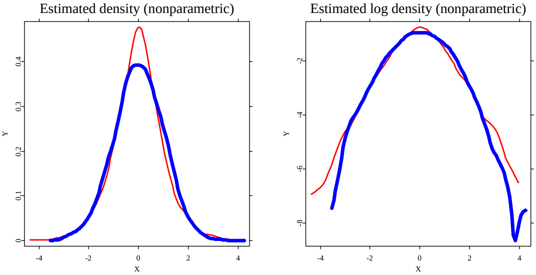


Figure 1: Densities (left) and log-densities (right) of the devolatilized return of daily DEM/USD FX rate from 1979/12/01 to 1994/04/01 (3720 observations). The kernel density estimate of the residuals (red) and the normal density (blue) with $\hat{h} = 0.55$ (rule of thumb). GHADAfx.xpl

A Stylized Fact in Financial Markets

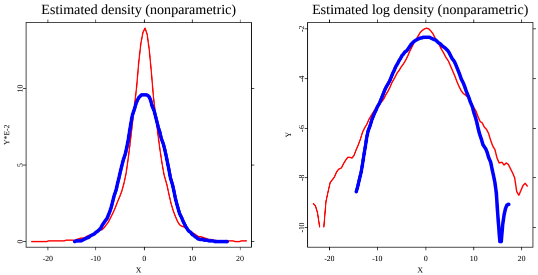
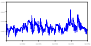


Figure 2: Densities (left) and log-densities (right) of the devolatilized DEM/USD return on the basis of GARCH(1,1) fit: $\hat{\sigma}_t^2 = 1.65e - 06 + 0.07 r_{t-1}^2 + 0.89 \sigma_{t-1}^2 + \varepsilon_t$. The kernel density estimate of the residuals (red) and the normal density (blue) with $\hat{h} = 2.17$ (rule of thumb). GHADAgarch.xpl



Risk Management Models

Heteroscedastic model

$$R_t = \sigma_t \varepsilon_t, \; t = 1, 2, \cdots$$

R_t (log) return, σ_t volatility, ε_t i.i.d. stochastic term.

Typical assumptions

1 The stochastic term is normally distributed, $\varepsilon_t \sim \mathrm{N}(0, 1)$.

2 A time-homogeneous structure of volatility:

- ARCH model, Engle (1995)
- GARCH model, Bollerslev (1995)
- Stochastic volatility model, Harvey, Ruiz and Shephard (1995)

Improvements

A The generalized hyperbolic (GH) distribution family fits the empirical distribution observed in financial markets.

- Hyperbolic (HYP) distribution in finance, Eberlein and Keller (1995),
- GH distribution + (parametric) stochastic volatility model, Eberlein, Kallsen and Kristen (2003).

B A time inhomogeneous model yields precise volatility estimation.

- Adaptive volatility estimation + normal distribution, Mercurio and Spokoiny (2004).

Combine A & B!

Motivation

GH distribution + adaptive volatility estimation (GHADA)

1. Adaptive volatility technique to estimate the volatility σ_t by $\hat{\sigma}_t$
2. Standardize the returns using
$$\varepsilon_t = \frac{R_t}{\hat{\sigma}_t}$$
3. Maximum likelihood to estimate the parameters $(\lambda, \alpha, \beta, \delta, \mu)^\top$ of GH distribution
4. Apply to risk measurement, e.g. Value at Risk (VaR) and TailVaR.

Backtesting: GHADA performs better than a model based on normal distribution.

Outline

1. Motivation ✓
2. Generalized hyperbolic (GH) distribution and its maximum likelihood (ML) estimation
3. Adaptive volatility estimation
4. Standardized (devolatilized) returns
5. VaR applications
6. Multivariate VaR and independent component analysis (ICA)

All calculations may be replicated in XploRe.

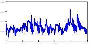
Generalized Hyperbolic (GH) Distribution

$X \sim GH$ with density:

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

Where $\iota^2 = \alpha^2 - \beta^2$, $K_{\lambda}(\cdot)$ is the modified Bessel function of the third kind with index λ: $K_{\lambda}(x) = \frac{1}{2}\int_0^{\infty} y^{\lambda-1} exp\{-\frac{x}{2}(y+y^{-1})\}\, dy$
Furthermore, the following conditions must be fulfilled:

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



GH parameters μ, δ

μ and δ control the location and the scale.

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

The term in the big brackets is location and scale invariant.

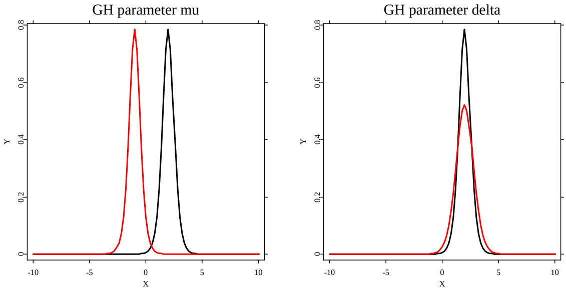
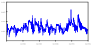


Figure 3: The GH pdf (black) with $\lambda = -0.5$, $\alpha = 3$, $\beta = 0$, $\delta = 1$ and $\mu = 2$. The left red line is obtained for $\mu = -3$ and the right red line is for $\delta = 2$ holding the other parameters constant. GHADAghexample.xpl



GH parameters β

β describes the skewness. For a symmetric distribution $\beta = 0$ according to the following lemma.

Lemma

The linear transformation $Y = aX + b$ of $X \sim GH$ is again GH distributed with parameters $\lambda_Y = \lambda$, $\alpha_Y = \alpha/|a|$, $\beta_Y = \beta/|a|$, $\delta_Y = \delta|a|$ and $\mu_Y = a\mu + b$.

$$f_{GH}(y = -x; \lambda, \alpha, \beta, \delta, -\mu) \overset{\beta=0}{=} f_{GH}(x; \lambda, \alpha, \beta, \delta, \mu)$$

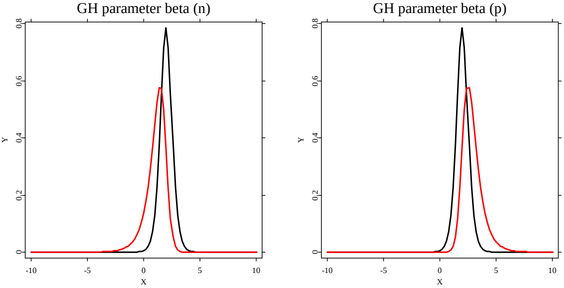


Figure 4: The GH pdf (black) with $\lambda = -0.5$, $\alpha = 3$, $\beta = 0$, $\delta = 2$ and $\mu = 2$. The left red line is obtained for $\mu = -3$ and the right red line is for $\delta = 2$ holding the other parameters constant. GHADAghexample.xpl

GH parameters α

α has an effect on kurtosis.

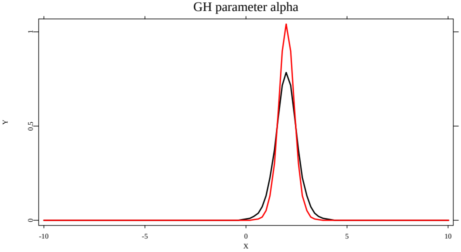


Figure 5: The GH pdf (black) with $\lambda = -0.5$, $\alpha = 6$, $\beta = 0$, $\delta = 1$ and $\mu = 2$. The left red line is obtained for $\mu = -3$ and the right red line is for $\delta = 2$ holding the other parameters constant. GHADAghexample.xpl

Subclass of GH distribution

The parameters $(\mu, \delta, \beta, \alpha)^\top$ *can be interpreted as trend, riskiness, asymmetry and the likeliness of extreme events.*

Hyperbolic (HYP) distributions: $\lambda = 1$,

$$f_{HYP}(x; \alpha, \beta, \delta, \mu) = \frac{\iota}{2\alpha\delta K_1(\delta\iota)} e^{\{-\alpha\sqrt{\delta^2+(x-\mu)^2}+\beta(x-\mu)\}}, \qquad (1)$$

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

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

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

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

Tail behavior of GH distribution

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

where $a(x) \sim b(x)$ as $x \to \infty$ means that both $a(x)/b(x)$ and $b(x)/a(x)$ are bounded as $x \to \infty$.

Comparison with other distributions:

- Normal distribution: $f_{Normal} = \frac{1}{\varsigma\sqrt{(2\pi)}} e^{-\frac{(x-\mu)^2}{2\varsigma^2}}$
- Laplace distribution: $f_{Laplace} = \frac{1}{2\varsigma} e^{-|x-\mu|/\varsigma}$
- Cauchy distribution: $f_{Cauchy} = \frac{1}{\varsigma\pi[1+(x-M)^2/\varsigma^2]}$

where μ is the location parameter, ς is the scale parameter and M is the Median.

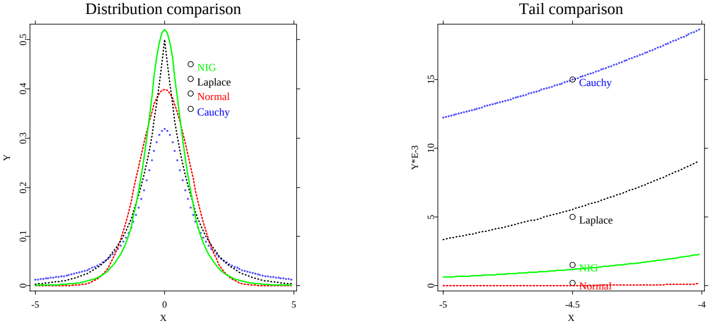
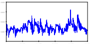


Figure 6: Graphical comparison of the NIG distribution (line), standard normal distribution (dashed), Laplace distribution (dotted) and Cauchy distribution (dots). GHADAtail.xpl

Maximum Likelihood (ML) Estimation

$$
\begin{aligned}
L_{HYP} &= n\log\iota - n\log 2 - n\log\alpha - n\log\delta - n\log K_1(\delta\iota) \\
&+ \sum_{t=1}^{n}\{-\alpha\sqrt{\delta^2+(x_t-\mu)^2}+\beta(x_t-\mu)\} \\
L_{NIG} &= n\log\alpha + n\log\delta - n\log\pi + n\delta\iota \\
&+ \sum_{t=1}^{n}\left[\log K_1\left\{\alpha\sqrt{\delta^2+(x_t-\mu)^2}\right\} - \frac{1}{2}\log\{\delta^2+(x_t-\mu)^2\}\right] \\
&+ \sum_{t=1}^{n}\beta(x_t-\mu)
\end{aligned}
$$



Example 1: DEM/USD exchange rate

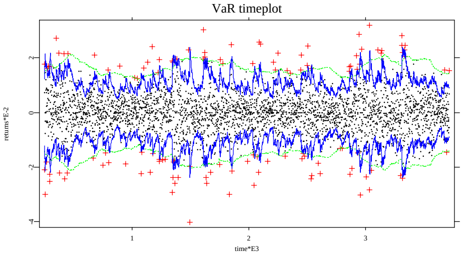
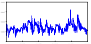


Figure 7: The time plot of VaR forecasts using EMA (green) and RMA (blue) and the associated changes (dots) of the P&L of the DEM/USD rates. Exceptions are marked in red. GHADArevartimeplot.xpl SFEVaRbank.xpl



ML estimators of HYP distribution: $\hat{\alpha} = 1.744$, $\hat{\beta} = -0.017$, $\hat{\delta} = 0.782$ and $\hat{\mu} = 0.012$.

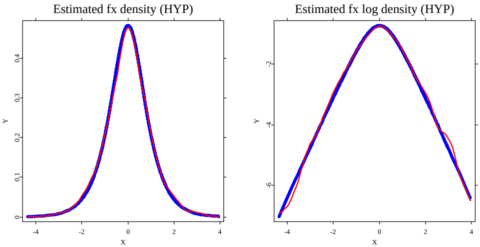


Figure 8: The estimated density (left) and log density (right) of the standardize returns of FX rates (red) with nonparametric kernel ($\hat{h} = 0.55$) and a simulated HYP density (blue) with the maximum likelihood estimators. GHADAfx.xpl

The HYP likelihood surface w.r.t. β and μ on the basis of DEM/USD data.

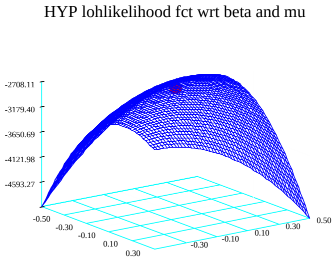
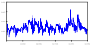


Figure 9: The partial HYP likelihood surface of the standardize returns of FX rates, the largest ML is marked in red. GHADAliksurf.xpl



ML estimators of NIG distribution: $\hat{\alpha} = 1.340$, $\hat{\beta} = -0.015$, $\hat{\delta} = 1.337$ and $\hat{\mu} = 0.010$.

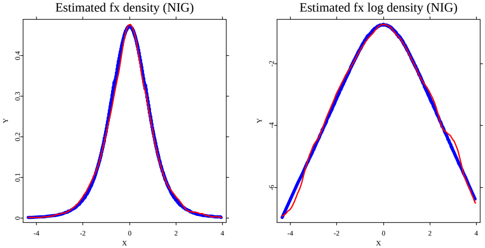
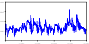


Figure 10: The estimated density (left) and log density (right) of the standardize return of FX rates (red) with nonparametric kernel ($\hat{h} = 0.55$) and a simulated NIG density (blue) with the maximum likelihood estimators. GHADAfx.xpl



Example 2: A German bank portfolio (kupfer.dat)

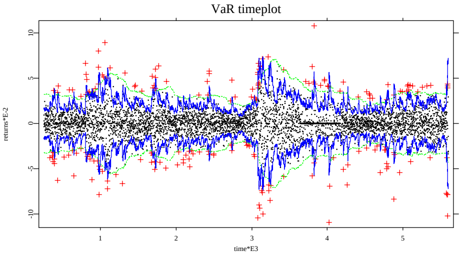


Figure 11: The time plot of VaR forecasts using EMA (green) and RMA (blue) and the associated changes (dots) of the P&L of the German bank portfolio. Exceptions are marked in red. GHADArevartimeplot.xpl

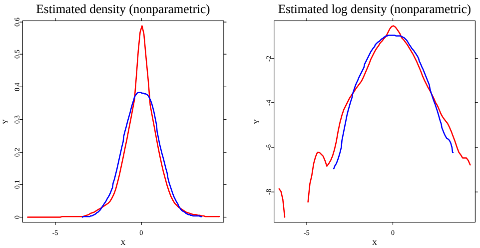
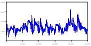


Figure 12: Graphical comparison of densities (left) and log-densities (right) of a German bank portfolio rate (5603 observations). The kernel density estimate of the standardized residuals (red) and the normal density (blue) with $\hat{h} = 0.61$ (rule of thumb). GHADAkupfer.xpl



ML estimators of HYP distribution: $\hat{\alpha} = 1.819$, $\hat{\beta} = -0.168$, $\hat{\delta} = 0.705$ and $\hat{\mu} = 0.145$.

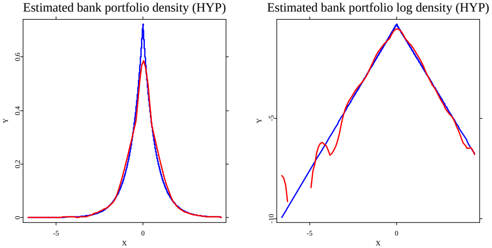


Figure 13: The estimated density (left) and log density (right) of the standardize return of bank portfolio rates (red) with nonparametric kernel ($\hat{h} = 0.61$) and a simulated HYP density (blue) with the maximum likelihood estimators. GHADAkupfer.xpl

ML estimators of NIG distribution: $\hat{\alpha} = 1.415$, $\hat{\beta} = -0.171$, $\hat{\delta} = 1.254$ and $\hat{\mu} = 0.146$.

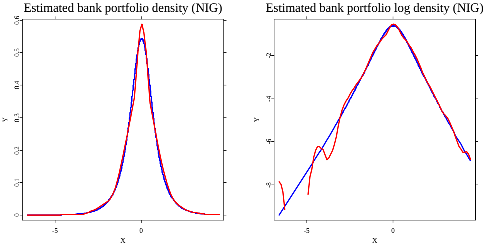


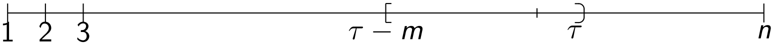
Figure 14: The estimated density (left) and log density (right) of the standardize return of bank portfolio rates (red) with nonparametric kernel ($\hat{h} = 0.61$) and a simulated NIG density (blue) with the maximum likelihood estimators. GHADAkupfer.xpl

Adaptive Volatility Estimation

Adaptive Volatility Estimation

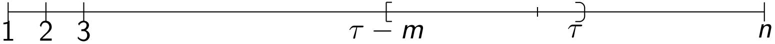
Assumption:

For a fixed point τ, volatility is locally time-homogeneous in a short time interval $[\tau - m, \tau)$, thus we can estimate $\hat{\sigma}_\tau^2 = \hat{\sigma}_I^2 = \frac{1}{|I|} \sum_{t \in I} R_t^2$, where $|I|$ is the number of observations in $I = [\tau - m, \tau)$.



Adaptive Volatility Estimation

Adaptive Volatility Estimation
For a fixed point τ, volatility is locally time-homogeneous in a short time interval $[\tau - m, \tau)$, thus we can estimate $\hat{\sigma}_\tau^2 = \hat{\sigma}_I^2 = \frac{1}{|I|}\sum_{t \in I} R_t^2$, where $|I|$ is the number of observations in $I = [\tau - m, \tau)$.



Questions:

- How to estimate the volatility? ✓
- How to specify the time homogeneous interval?

Volatility estimation

Volatility estimation

Power transformation yields lighter tails. The random variable $|R_t|^\gamma$ is distributed more evenly.

For every $\gamma > 0$, we have

$$\begin{aligned}
\mathsf{E}(|R_t|^\gamma|\mathcal{F}_{t-1}) &= \sigma_t^\gamma \mathsf{E}(|\varepsilon_t|^\gamma|\mathcal{F}_{t-1}) = C_\gamma \sigma_t^\gamma \\
\mathsf{E}[(|R_t|^\gamma - C_\gamma\sigma_t^\gamma)^2|\mathcal{F}_{t-1}] &= \sigma_t^{2\gamma}\mathsf{E}[(|\varepsilon_t|^\gamma - C_\gamma)^2|\mathcal{F}_{t-1}] \\
&= \sigma_t^{2\gamma} D_\gamma^2
\end{aligned}$$

$$|R_t|^\gamma = C_\gamma\sigma_t^\gamma + D_\gamma\sigma_t^\gamma\zeta_t, \tag{4}$$

where C_γ is the conditional mean and D_γ^2 the conditional variance of $|\varepsilon|^\gamma$ and $\zeta_t = (|\varepsilon_t|^\gamma - C_\gamma)/D_\gamma$ is i.i.d. with mean 0.

Denote by $\theta_t = C_\gamma \sigma_t^\gamma$ the **conditional mean** of $|R_t|^\gamma$.
In a time-homogeneous interval I, the constant $\theta_I = C_\gamma \sigma_I^\gamma$ can be estimated by $\hat{\theta}_I$:

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

We employ the "nearly constant" θ_I to determine the length of the interval I.

Use $|R_t|^\gamma = \theta_t + D_\gamma \sigma_t^\gamma \zeta_t$:

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

$$\mathrm{Var}[\hat{\theta}_I | \mathcal{F}_t] = \frac{s_\gamma^2}{|I|^2} \, \mathrm{E} \sum_{t \in I} \theta_t^2$$

where $s_\gamma = D_\gamma / C_\gamma$. We denote $v_I^2 = \frac{s_\gamma^2}{|I|^2} \, \mathrm{E} \sum_{t \in I} \theta_t^2$, the conditional variance of $\hat{\theta}_I$, whose estimator is

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

Time homogeneous interval

Split interval I into $J \subset I$ and $I \backslash J \subset I$.

$$\overset{[}{\tau - m} \qquad I \backslash J \qquad | \qquad J \qquad \overset{]}{\tau - 1}$$

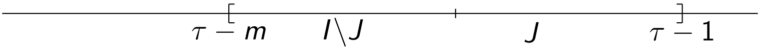
Denote $\hat{\theta}_{I \backslash J}$ and $\hat{\theta}_J$ as the estimators of the subintervals in the time-homogeneous interval I. Then the deviation $\Delta = \hat{\theta}_{I \backslash J} - \hat{\theta}_J$ must be small.

$$|\hat{\theta}_{I \backslash J} - \hat{\theta}_J| \quad \leq \quad T_{I,\tau}. \tag{5}$$

where $T_{I,\tau}$ is an unknown critical value in the homogeneity test.

$$I^* = \max \{I : I \text{ fulfills } (7)\}.$$

Time homogeneous interval



$$|\hat{\theta}_{I\backslash J} - \hat{\theta}_J| \quad \leq \quad T_{I,\tau}.$$

Questions:

- How to estimate the volatility? ✓
- How to specify the time homogeneous interval? ✓
- How to choose γ?
- How to specify $T_{I,\tau}$?

Homogeneity test

Lemma

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

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

where $\zeta_\gamma = (|\varepsilon_t|^\gamma - C_\gamma)/D_\gamma$ *and* ε *is a GH distributed stochastic term.*

If Lemma 2 holds, then $\Upsilon_t = \exp\left(\sum_{s=1}^{t} p_s \zeta_s - (a_\gamma/2) \sum_{s=1}^{t} p_s^2\right)$ is a supermartingale, where p_s is a predictable process w.r.t. the information set $\mathcal{F}_{t-1}$. Since:

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

i.e. $\mathsf{E}(\Upsilon_t|\mathcal{F}_{t-1}) \leq \Upsilon_{t-1}$.

Theorem

If $R_1, ..., R_\tau$ obey the heteroscedastic model and the volatility coefficient σ_t satisfies the condition $b \le \sigma_t^2 \le bB$ with some positive constant b and B, then given a large value η it holds for the estimate $\hat{\theta}_I$ of θ_τ :

$$
\begin{aligned}
& \mathrm{P}(|\hat{\theta}_I - \theta_\tau| > \Delta_I(1 + \eta s_\gamma |I|^{-1/2}) + \eta \hat{v}_I) \\
\le \quad & 4\sqrt{e}\eta(1 + logB)\exp(-\tfrac{\eta^2}{2a_\gamma(1+\eta s_\gamma |I|^{-1/2})^2}).
\end{aligned}
$$

where Δ_I is the 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 $\eta\hat{v}_I(\approx T_{I,\tau})$ for $\tau \in I$ with a high probability, since in a time homogeneous interval the squared bias Δ_I is negligible.

Test homogeneity: Under homogeneity $|\hat{\theta}_I - \theta_\tau|$ is bounded by $\eta \hat{v}_I$ provided that η is sufficiently large.

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

Based on the triangle inequality, we get: $\hat{\theta}_{I \setminus J} - \hat{\theta}_J$ is bounded by $\eta(\hat{v}_{I \setminus J} + \hat{v}_J)$ for $J \subset I$, i.e.

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

where $\eta' = \eta s_\gamma$.

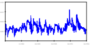
Test homogeneity: Under homogeneity $|\hat{\theta}_I - \theta_\tau|$ is bounded by $\eta \hat{v}_I$ provided that η is sufficiently large.

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

$$|\hat{\theta}_{I\setminus J} - \hat{\theta}_J| \le T_{I,\tau} = \eta(\hat{v}_{I\setminus J} + \hat{v}_J) = \eta'(\sqrt{\hat{\theta}_J^2 |J|^{-1}} + \sqrt{\hat{\theta}_{I\setminus J}^2 |I\setminus J|^{-1}}),$$

Questions:

- How to estimate the volatility? ✓
- How to specify the time homogeneous interval? ✓
- How to specify $T_{I,\tau}$? ✓
- How to choose γ?
- How to choose η'?



Cross-validation (CV) method:

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

where t_0 is the starting point.
Choice of transformation parameter γ: we choose $\gamma = 0.5$ to compare with the normal distribution based model.

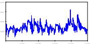
Cross-validation (CV) method:

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

Choice of transformation parameter γ: we choose $\gamma = 0.5$ to compare with the normal distribution based model.

Questions:

- How to estimate the volatility? ✓
- How to specify the time homogeneous interval? ✓
- How to choose γ? ✓
- How to choose η'? ✓

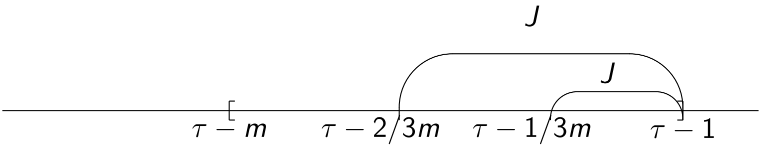


Iteration

Start from a short homogeneous interval $[\tau - m_0, \tau - 1]$, the algorithm consists of 4 steps.

- Step 1: At $\tau - 1$, enlarge the interval I from $[\tau - m_0, \tau)$ to $[\tau - k \times m_0, \tau)$, i.e. $m = k \times m_0$. The parameters m_0 and k are integers specified according to data. Values of $m_0 = 5$ and $k = 2$ are recommended.

- Step 2: Inside interval I, do multiple homogeneity tests based on subintervals $J = [\tau - \frac{1}{3}m, \tau)$ until $J = [\tau - \frac{2}{3}m, \tau)$.



- Step 3: If homogeneity hypothesis is rejected at point s, the loop stops. Otherwise go back to Step 1.
- Step 4: Do Step 1 to Step 3 for $t \in [t_0, \tau - 1]$ with different η's. Choose η' that gives the minimal global forecast error.

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

Simulation

Goal: Estimate the local volatility using GHADA.
Simulation: 200 processes with HYP and NIG distribution with $(\alpha, \beta, \delta, \mu)^\top = (2, 0, 1, 0)^\top$ respectively. Each process has $T = 1000$ observations.
Parameters: starting point $t_0 = 201$, power transformation parameter $\gamma = 0.5$, $m_0 = 5$ and $k = 2$.

Case 1: 200 simulated HYP random variables and

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

Case 2: 200 simulated NIG random variables

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

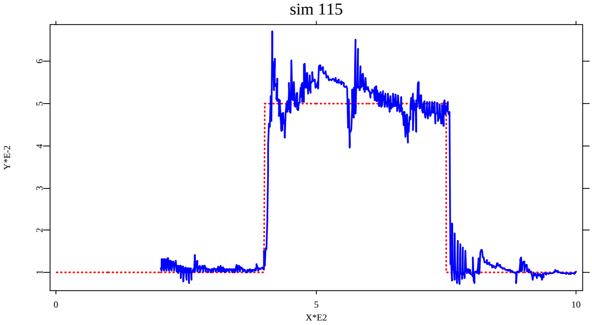


Figure 15: The estimated local volatilities for simulation 115 (HYP).
GHADAsim1.xpl

http://ise.wiwi.hu-berlin.deychen/ghada/simulation1.AVI

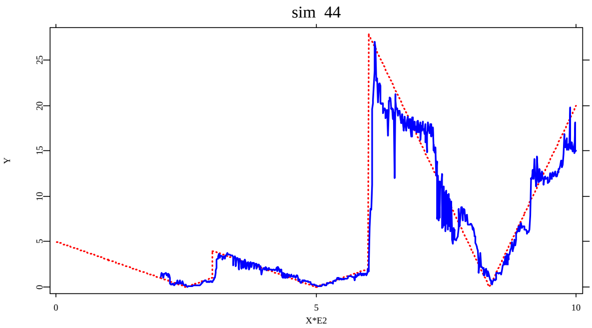


Figure 16: The estimated local volatilities for simulation 44 (NIG).
GHADAsim2.xpl

http://ise.wiwi.hu-berlin.deychen/ghada/simulation2.AVI

Sensitivity Analysis: Define a percentage rule that tells us after how many steps a sudden jump is detected at 40%, 50% or 60% of the jump level.

	mean	std	max	min
σ_1: Detection delay to the first jump at $t = 400$				
40% rule	5.9	2.4	15	1
50% rule	6.9	2.6	19	2
60% rule	7.9	2.9	19	2
σ_1: Detection delay to the second jump at $t = 750$				
40% rule	11.8	4.4	39	3
50% rule	13.5	6.5	58	5
60% rule	15.9	10.9	98	6

Table 1: Descriptive statistics for the detection delay of the sudden vola jumps.

	mean	std	max	min
σ_2: Detection delay to the first jump at $t = 300$				
40% rule	4.9	2.4	13	0
50% rule	6.2	3.0	18	2
60% rule	7.6	4.2	33	2
σ_2: Detection delay to the first jump at $t = 600$				
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 2: (Continued) Descriptive statistics for the detection delay of the sudden vola jumps.

Data set and devolatilized return

	DEM/USD	bank portfolio
period	791201 to 940401	
observations	3720	5603
mean	-0.0052	0.0113
std	0.9938	0.9264
skewness	-0.0121	-0.0815
kurtosis	4.0329	5.1873

Table 3: Descriptive statistics for the standardized residuals of DEM/USD data and bank portfolio data. The data sets are available in http://www.quantlet.org/mdbase/.

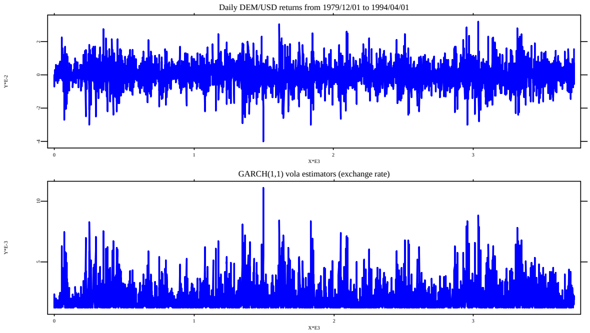


Figure 17: The return process of DEM/USD exchange rates (top), the GARCH(1,1) ($\hat{\sigma}_t^2 = 1.65e - 06 + 0.07 r_{t-1}^2 + 0.89 \sigma_{t-1}^2 + \varepsilon_t$) volatility estimates (bottom).
GHADAgarch.xpl

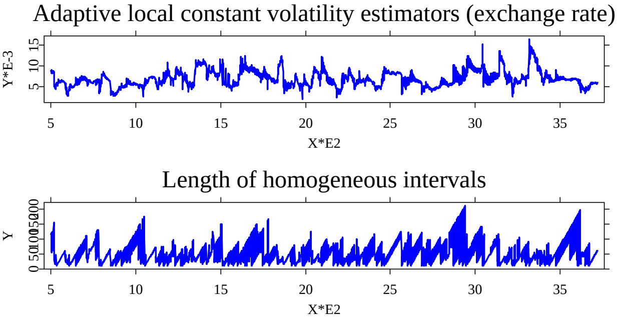


Figure 18: The adaptive volatility estimates (top) and the lengths of the homogeneous intervals for $t_0 = 501$, $\eta' = 1.06$ and $m_0 = 5$. The average length of time homogeneous interval is 51. GHADAfx.xpl

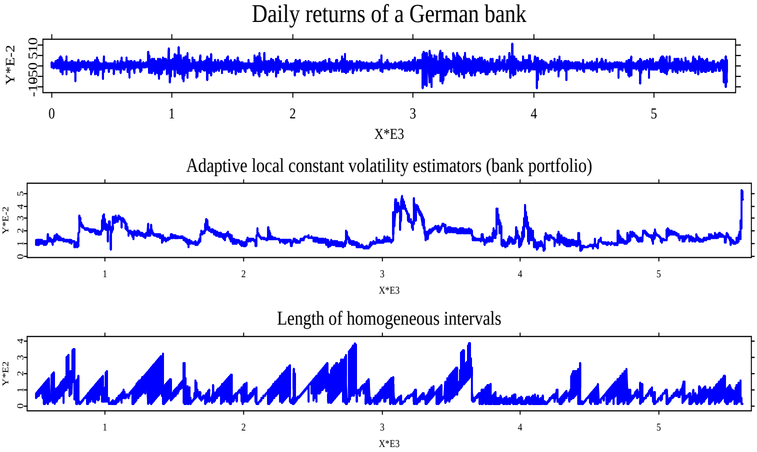
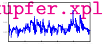


Figure 19: The return process of a German bank's portfolio (upper), its adaptive volatility estimates (middle) for $t_0 = 501$, $\eta' = 1.23$ and $m_0 = 5$, the lengths of the homogeneous intervals (bottom). The average length of time homogeneous interval is 72.

GHADAkupfer.xpl



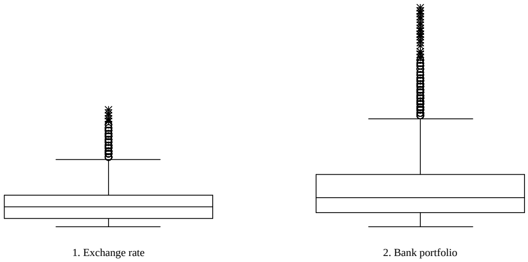


Figure 20: Boxplots of the DEM/USD exchange rates (left) and the German bank portfolio data (right). The mean values of the homogeneous interval length are 51 for DEM/USD and 72 for bank portfolio data. GHADAboxplot.xpl

Value at Risk (VaR)

q_p is the p-th quantile of the distribution of ε_t, i.e.
$P(\varepsilon_t < q_p) = p$.

$$P(R_t < \sigma_t q_p \,|\, \mathcal{F}_{t-1}) = p$$

$$\mathrm{VaR}_{p,t} = \sigma_t q_p$$

$\hat{\sigma}_t$ are estimated by the described adaptive procedure.
q_p is given by the quantile of the HYP or NIG distribution

GHADA VARs

Parameters estimation is based on the previous 500 observations (standardized returns), which varies little.

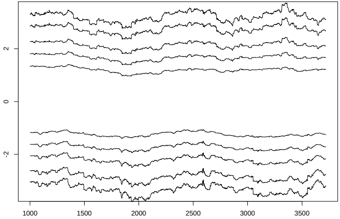
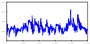


Figure 21: Quantiles based on DEM/USD data vary over time. From the top the evolving HYP quantiles for $p = 0.995$, $p = 0.99$, $p = 0.975$, $p = 0.95$, $p = 0.90$, $p = 0.10$, $p = 0.05$, $p = 0.025$, $p = 0.01$, $p = 0.005$.



GHADA model vs. normal model

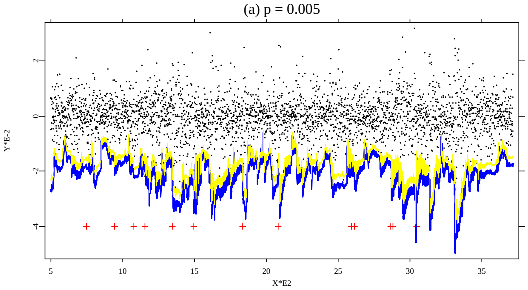


Figure 22: Value at Risk forecast plots for DEM/USD data. (a) $p = 0.005$. Dots denote the exchange rate returns. Exceptions relative to the HYP quantile are displayed as $+$. The blue line is the VaR forecasts based on GHADA while those based on the normality is colored in yellow. GHADAfxvar.xpl

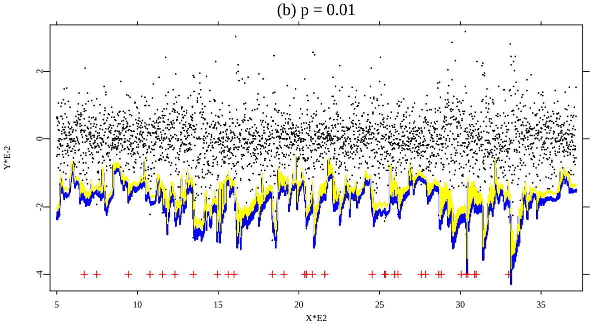
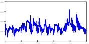


Figure 23: Value at Risk forecast plots for DEM/USD data. (b) $p = 0.01$.Dots denote the exchange rate returns. Exceptions relative to the HYP quantile are displayed as $+$. The blue line is the VaR forecasts based on GHADA while those based on the normality is colored in yellow. GHADAfxvar.xpl



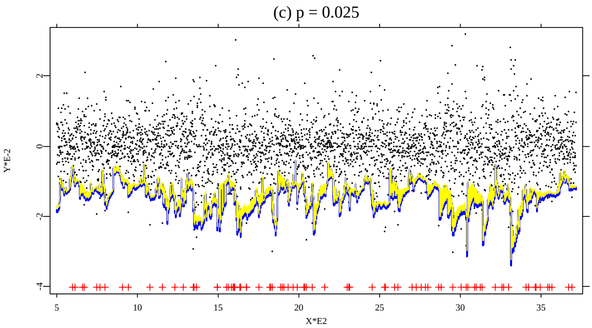
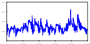


Figure 24: Value at Risk forecast plots for DEM/USD data. (c) $p = 0.025$.Dots denote the exchange rate returns. Exceptions relative to the HYP quantile are displayed as $+$. The blue line is the VaR forecasts based on GHADA while those based on the normality is colored in yellow. GHADAfxvar.xpl



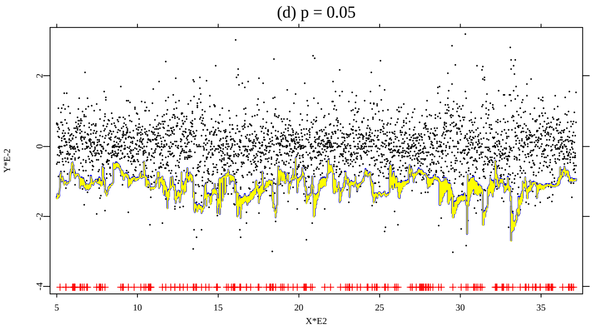
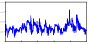


Figure 25: Value at Risk forecast plots for DEM/USD data. (d)$p = 0.05$.Dots denote the exchange rate returns. Exceptions relative to the HYP quantile are displayed as $+$. The blue line is the VaR forecasts based on GHADA while those based on the normality is colored in yellow. GHADAfxvar.xpl



Backtesting VaR

Testing VaR levels:

$$H_0 : \mathsf{E}\, N = pT \quad \text{vs.} \quad H_1 : \text{not } H_0 \tag{6}$$

where N is the number of the exceptions on the basis of T observations. Likelihood ratio statistic:

$$\mathrm{LR1} = -2 \log \left\{ (1-p)^{T-N} p^N \right\} + 2 \log \left\{ (1 - N/T)^{T-N} (N/T)^N \right\},$$

where $\mathrm{LR1}$ is asymptotically $\chi^2(1)$ distributed

Let I_t denote the indicator of exceptions at time point t, $t = 1, ..., T$, $\pi_{ij} = \mathrm{P}(I_t = j | I_{t-1} = i)$ be the transition probability with $i, j = 0$ or 1, and $n_{ij} = \sum_{t=1}^{T} I(I_t = j | I_{t-1} = i), i, j = 0, 1$.

Testing Independence:

$$H_0 : \pi_{00} = \pi_{10} = \pi, \pi_{01} = \pi_{11} = 1 - \pi \quad \text{vs.} \quad H_1 : \text{ not } H_0$$

Likelihood ratio statistic:

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

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, $\mathrm{LR2}$ is asymptotically $\chi^2(1)$ distributed as well.

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 4: Backtesting results for DEM/USD example. * indicates the rejection.

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 5: Backtesting results for bank portfolio example.

Multivariate VaR

Background:

- Multivariate GH distribution, Prause(1999), Schmidt, Hrycej and Stützle(2003).
- First two principal components and multinormal distribution, Härdle, Herwatz and Spokoiny(2003).

New idea: Independent Component Analysis (ICA) + univariate GHADA

ICA

Definition of ICA:

- Observation vector $x_t = (x_{1,t}, \cdots, x_{d,t})^\top$
- Independent vector $s_t = (s_{1,t}, \cdots, s_{d,t})^\top$: assumption in ICA - the components s_i are statistically independent, $i = 1, \cdots, d$.
- Unknown mixing matrix $A - d \times d$ dimensions: $x = As$

Example: Two independent components s_1 and s_2 are uniform distributed. The mixing matrix $A = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$.

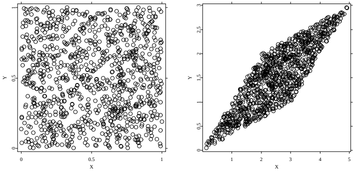


Figure 26: The joint distribution of the independent components s_1 and s_2. Below is the joint distribution of the observed mixtures x_1 and x_2. GHADAicaexample.xpl

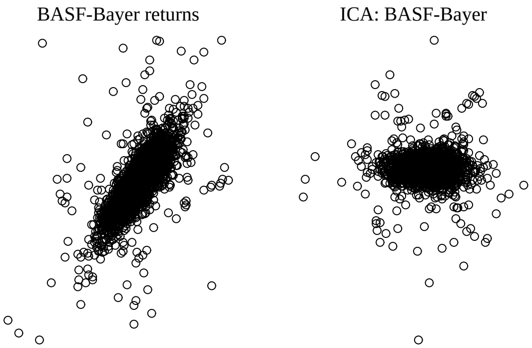


Figure 27: Independent component analysis on the basis of Allianz and Bayer returns from 1974-01-02 to 1996-12-30. GHADAica.xpl

Conclusion

- The adaptive volatility estimation method by Mercurio and Spokoiny (2004) is applicable to a general model with generalized hyperbolic innovations.
- The critical value can be chosen by cross-validation method.
- 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 innovations can be perfectly modelled by the class of generalized hyperbolic distributions.

- The proposed approach can be applied easily to risk measures such as value at risk, expected shortfall, and so on.
- We have got a justification of the proposed appoach for use in risk management by backtestings of value at risk model applied to real data.

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.

Bibby, B. M. and Sørensen, M. (2001). Hyperbolic Processes in Finance, *Technical Report 88*, University of Aarhus, Aarhus School of Business.

Bollerslev, T. (1995). Generalized autoregressive conditional heteroskedasticity, ARCH, selected readings, *Oxford University Press.*

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, selected readings, *Oxford University Press.*

Härdle, W., Herwartz, H. and Spokoiny, V. (2003). Time Inhomogeneous Multiple Volatility Modelling, *Journal of Financial Econometrics* **1**, 55–95.

Härdle, W., Kleinow, T. and Stahl, G. (2002). Applied Quantitative Finance, *Springer-Verlag Berlin Heidelberg New York*.

Harvey, A., Ruiz, E. and Shephard N. (1995). Multivariate stochastic variance models, ARCH, selected readings, *Oxford University Press.*

Hyvärinen, A., Karhunen, J. and Oja, E. (2001). Independet Component Analysis, *John Willey & Sons, INC.*

Jorion, P. (2001). Value at Risk, *McGraw-Hill*.

Liptser, R. and Spokoiny, V. (2000). Deviation Probability Bound for Martingales with Applications to Statistical Estimation, *Statistics and Probability Letters* **46**, 347–357.

Mercurio, D. and Spokoiny, V. (2004). Statistical Inference for Time Inhomogeneous Volatility Models, *Annals of Statistics* **32**, 577–602.

Mercurio, D. and Spokoiny, V. (2004). Volatility Estimation via Local Change Point Analysis with Applications to Value-at-Risk, *working paper*.

Prause, K. (1999). The Generalized Hyperbolic Model: Estimation, Financial Derivatives and Risk Measures, *dissertation*.

Schmidt, R., Hrycej, T. and Stützle, E. (2003). Multivariate Distribution Models with Generalized Hyperbolic Margins, *working paper*.