Strategic Long-Term Financial Risks
The One-Dimensional Case

Roger Kaufmann
Pierre Patie
RiskLab
D-MATH
ETH-Zentrum
CH-8092 Zürich
Switzerland
kaufmann@ / patie@math.ethz.ch

December 2, 2003

Abstract. The development of a methodology that could be used for the measurement of strategic long-term financial risks is becoming an important task. Existing modelling instruments allow for a good measurement of market risks of trading books over relatively small time intervals. However, these approaches might have some severe deficiencies if they are applied to longer time periods. In this paper we give an overview on methodologies that are proposed to model the evolution of risk factors over a long horizon. We investigate in detail the statistical properties and the behaviour of financial time series at different frequencies. Then, we test the different models on these data by backtesting expected shortfall predictions.

Contents

Key words and phrases. Value-at-risk, expected shortfall, random walk models, vector auto regressive model, jump-diffusion models, stochastic volatility models, extreme value theory, scaling rules.

Research of both authors was supported by Credit Suisse Group, Swiss Re and UBS AG through RiskLab, Switzerland.

1 Introduction

Most of the current research in asset management and market risk management focuses on short-term risk (daily, weekly). Clearly, in many cases, a long-term analysis (e.g. quarterly, yearly) is just as relevant and important. Indeed, in financial institutions risk measures are typically calculated for a one-day to two-week horizon. In most other institutions, by contrast, assets and liabilities can be priced only occasionally, and risk exposures must be measured and managed over these much longer time spans. While for the measurement of short-term financial risks some consistent and reliable frameworks already exist, there has not been much work done on longer time horizons. Only relatively few papers on this topic can be found in the academic literature.

Existing modelling instruments such as RiskMetrics allow for a relatively good measurement of market risks of trading books. These models, however, have some severe deficiencies if they are applied to longer time periods (typically one year), as needed in the case of strategic investments of institutional investors or within insurance. Recently, some results are available, like for instance the ForeSight Technical Document [27] of the RiskMetrics Group which focuses on forecasting beyond a two year horizon.

In this paper we develop a theoretically well-understood and empirically-founded conceptual framework for the measurement of long-term financial risk of strategic investment portfolios. Our criterion to judge the appropriateness of different models is the reliability of expected shortfall estimates of profit and loss distributions for financial portfolios over a period of one year. We choose expected shortfall as risk measure because it is coherent in the sense of Artzner et al. [3].

An obvious approach one may take to model yearly returns is to extend methodologies from short-term risks to longer time horizons. It is well known that daily (or higher frequency) returns are dependent due to volatility clustering, while returns over a longer time horizon (fortnightly, monthly data) are closer to being independent. Because of the weaker dependence for lower frequency data, some time aggregation rules may perform better than in the short-term case. On the other hand we also have to take into account the serious statistical restrictions due to the small size of low frequency data. Therefore the pivotal point is the choice of the data frequency on which we calibrate our models. There is a tradeoff between bias and variance.

In this paper we present our findings on the evolution of single risk factors. Sections 3 to 6 give an overview on approaches that we will use to model the risk factors. In these sections we first investigate dynamical models like random walks, AR(p) and GARCH(1,1) models, which allow to model price changes. Then we propose a static approach based on heavy tailed distributions. We point out that we follow the same plan for presenting each approach. First, we describe the model including the statistical tools used to estimate the parameters, then we show how to aggregate returns. Finally, we give the methodology used to compute the risk measures. In Section 7 we present some characteristics of data sets which are representative for a portfolio. In Section 8 we compare the different models by backtesting expected shortfall predictions on exchange rate data, stock indices and 10-year government bonds. We also provide a variance analysis for the time series models, and we compute confidence intervals for expected shortfall and value-at-risk for the random walk approach. The final section contains the conclusions that can be drawn from our investigations.

2 Definitions and notation

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space on which all random variables are assumed to be defined. We recall some useful definitions which will be the basis of our study. Following Artzner et al. [3], we define:

Definition 2.1 Let $\mathcal{G}$ be the set of all risks. Here, we restrict ourselves to financial risks. A measure of risk is a mapping from $\mathcal{G}$ into $\mathbb{R}$. The requirements of elements of $\mathcal{G}$ depend on the risk measure considered.

Definition 2.2 Given $p \in (0,1)$, the *value-at-risk* at a level p, denoted by VaR^p, of the return R is given by

$$\mathrm{VaR}^p(R) = -\inf\{x \in \mathbb{R} \mid \mathbb{P}[R \leq x] \geq p\},$$

i.e. VaR^p is the negative p-quantile of R.

Definition 2.3 The *expected shortfall* at a level p, denoted by ES^p, is defined by

$$\mathrm{ES}^p(R) = -\mathbb{E}[R \mid R < -\mathrm{VaR}^p(R)], \quad R \in L^1(\mathbb{P}).$$

The main difference between these two risk measures is the failure of value-at-risk in general to fulfil the subadditivity property required for a risk measure to be coherent.

Definition 2.4 A measure of risk ρ is *subadditive*, if for all $R^a, R^b \in \mathcal{G}$

$$\rho(R^a + R^b) \leq \rho(R^a) + \rho(R^b).$$

Artzner et al. [3] explain why the subadditivity property is a natural requirement and show that this property holds under certain restrictions on the discounted risks, as well as on the underlying probability space. Rockafellar and Uryasev [34] give an equivalent representation of the conditional expected value as the solution of an optimization problem only requiring the probability distribution of the risk to be continuous. From this minimization problem, Pflug [32] shows the convexity (including the subadditivity property) of the expected shortfall. Finally, Acerbi and Tasche [1] give a definition of the expected shortfall for which its coherence property holds even in the case of discontinuous underlying loss distributions.

Notation. Throughout the rest of the paper we use the following notation:

- S_t: asset price at time t, it is assumed that for each $t \geq 0$, S_t is $\mathcal{F}_t-$adapted,
- $s_t = \log S_t$: asset log-price,
- h: denotes the length (in days) of 1 period,
- $R_t = \frac{S_t - S_{t-h}}{S_{t-h}}$: 1-period (i.e. h-day) asset return,
- $r_t = s_t - s_{t-h}$: 1-period asset log-return,
- r_t^k: k-period (i.e. kh-day) asset log-return (think of kh days = 1 year),
- μ: mean of 1-period asset log-returns, assumed to be constant over time,
- $\overline{r}_t = r_t - \mu$: centered 1-period asset log-return.

3 Random walks

A straightforward way to model log-returns at different time horizons is the random walk approach with normal innovations. As this approach concentrates on modelling the "non-extreme cases", it is a priori not clear whether this model is appropriate for estimating tails of distributions of log-returns. Nevertheless, focusing on long term risk measures, we will see that this model can yield useful insight into the behaviour of log-returns at long horizons. The first model we investigate is a random walk with a trend which is assumed to be constant over time. In other words, the 1-period log-returns r are assumed to be independent for non-overlapping periods and normally distributed with mean μ and variance σ^2. That is,

$$r_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2) \qquad \text{for } t \in h\mathbb{N}. \tag{3.1}$$

The mean of a sample of n 1-period log-returns can be estimated by

$$\widehat{\mu} = \frac{1}{n} \sum_{i=1}^{n} r_{ih},$$

and we use the following unbiased estimator for the 1-period variance:

$$\widehat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (r_{ih} - \widehat{\mu})^2.$$

3.1 Scaling rule.

When changes in the log-price are independent and identically distributed, which is the case when the log-price at time t can be written as follows

$$s_t = s_{t-h} + \epsilon_t,$$

where (ϵ_t) are i.i.d. with marginal distribution F_ϵ, then the 1-period log-returns have the same law as ϵ_t, that is $r_t \overset{\text{i.i.d.}}{\sim} F_\epsilon$.

Similarly the k-period log-returns can be written as

$$r_t^k = \sum_{i=0}^{k-1} \epsilon_{t-ih}.$$

For $\mathbb{E}[\epsilon_t] = 0$ and $\mathbb{E}[\epsilon_t^2] = \sigma^2 < \infty$, these k-period log-returns r_t^k have mean zero and standard deviation $\sqrt{k}\sigma$. Hence the 1-period volatility can be multiplied by the square-root-of-time to calculate the k-period volatility in the i.i.d. case. Under the further assumptions (3.1), we have the same scaling law for the quantiles.

It is well known that daily log-returns show neither independent nor normally distributed behaviour (see Fama [20], or Campbell et al. [8]). As we will observe in the empirical studies in Section 7, lower frequency data are closer to the normality assumption which is needed to apply the square-root-of-time rule (see also Campbell et al. [8] for similar investigations). Hence, this scaling rule should work better for models fitted to monthly or even lower frequency data.

In general, the square-root-of-time scaling rule cannot be applied directly to log-returns. We first have to subtract the mean. For instance, for i.i.d. normally distributed

1-period log-returns with constant mean $\mu = \mathbb{E}[\epsilon_t]$ and constant variance $\sigma^2 = \mathbb{E}[\epsilon_t^2]$, the following rule holds:

$$r_t^k \overset{\mathrm{d}}{\sim} k\mu + \sqrt{k}(r_t - \mu).$$

3.2 Estimation of value-at-risk and expected shortfall.

The square-root-of-time rule leads to

$$\widehat{\sigma}^k = \sqrt{k}\,\widehat{\sigma},$$

for the k-period standard deviation. The trend for k periods can be estimated by

$$\widehat{\mu}^k = k\,\widehat{\mu}.$$

This leads to the following formula for a k-period (think of 1 year) value-at-risk estimate at level p:

$$\widehat{\mathrm{VaR}}^p = -\Big(\exp\big(\widehat{\mu}^k + \widehat{\sigma}^k x^p\big) - 1\Big), \tag{3.2}$$

and the corresponding expected shortfall can be estimated by

$$\widehat{\mathrm{ES}}^p = -\Big(\exp\big(\widehat{\mu}^k + \frac{(\widehat{\sigma}^k)^2}{2}\big)\frac{\Phi(x^p - \widehat{\sigma}^k)}{p} - 1\Big), \tag{3.3}$$

where

x^p : p-quantile of a standard normal random variable,

Φ : cumulative standard normal distribution function.

3.3 Generalisation.

Alternatively to the previous approach, one of the following models which are based on current market information can be used:

a) Random walk with zero expected log-return (RWZ). This approach is often used for short-period forecasts.
b) Random walk with expected log-return equal to the forward premium (RWF). In this model the so-called forward premium is used for estimating a drift.

These two models are the ones proposed in Kim et al. [26]. We present them here in detail.

a) Random walk with zero expected log-return.

This first model is defined by two assumptions:

Assumption 1. Expected log-returns are equal to zero:

$$\mathbb{E}[r_t] = 0 \qquad \forall t.$$

Assumption 2. log-returns are independent, normally distributed with standard deviation σ.

Assumptions 1 and 2 imply that the logarithmic asset price follows a random walk with zero drift:

$$s_t - s_{t-h} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2), \quad t \in h\mathbb{N}.$$

For this model, for a sample of size n, the standard deviation σ^2 can be estimated by

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(s_{ih} - s_{(i-1)h})^2.$$

In model a) expected log-returns are zero. A generalisation of this is model b), which allows for interest rates and cash payouts.

b) Random walk with expected log-return equal to the forward premium.

A refinement of the RWZ model is the random walk with expected log-return equal to the forward premium (RWF). Again, two assumptions are needed to define the model:
Assumption 1'. Expected log-returns are equal to the so-called forward premium

$$\mathbb{E}[s_{t+h} \mid \mathcal{F}_t] - s_t = r_{t,t+h} - d_{t,t+h},$$

where $r_{t,t+h}$ denotes the risk free continuously compounded interest rate for the period $[t, t+h]$, and $d_{t,t+h}$ is the asset's 1-period (h-day) cash yield, expressed in currency units. The cash yield $d_{t,t+h}$ is $\mathcal{F}_t$-measurable. The asset's 1-period cash payouts at time $t+h$ can be written as $S_{t+h}(e^{d_{t,t+h}} - 1)$.

The idea underlying Assumption 1' is that the expected logarithm of the asset's discounted value at time $t+h$ (i.e. the expected logarithm of $S_{t+h}\, e^{d_{t,t+h}}\, e^{-r_{t,t+h}}$) should be equal to the logarithm of the asset's value at time t. Assumption 1' can be rewritten in the following way:

$$\mathbb{E}[\log(S_{t+h}\, e^{d_{t,t+h}}\, e^{-r_{t,t+h}}) \mid \mathcal{F}_t] = \log(S_t).$$

Assumption 2. log-returns are independent, normally distributed with standard deviation σ in each period $[t, t+h]$.

Assumptions 1' and 2 imply that the logarithmic asset price follows a weak version of a random walk. This means that increments $s_{t+h} - s_t$ are independent normally, but not identically distributed, since the forward premium $r_{t,t+h} - d_{t,t+h}$ fluctuates:

$$s_{t+h} - s_t - (r_{t,t+h} - d_{t,t+h}) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2).$$

The drift $r_{t,t+h} - d_{t,t+h}$ is random and time-varying, but $\mathcal{F}_t$-measurable.

In this model the asset prices and forward prices can be used to estimate σ. If forward price data are available for n time steps back, σ^2 can be estimated by

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(s_{ih} - f_{(i-1)h,ih})^2,$$

where $f_{t,t+h}$ is the logarithmic forward price:

$$\begin{aligned} f_{t,t+h} &= \log(S_t\, e^{r_{t,t+h} - d_{t,t+h}}) \\ &= s_t + r_{t,t+h} - d_{t,t+h}. \end{aligned}$$

If forward price data are only available for the more recent past, σ^2 can instead be estimated by

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(s_{ih} - s_{(i-1)h})^2.$$

Since the drift is neglected in this formula, this estimator is biased.

We generalise now from 1-period forecasts to k-period forecasts. The difference between the k-period ahead log-price s_{t+kh} and the current log-price s_t is modelled by a normal distribution with drift equal to the forward premium:

$$s_{t+kh} - s_t - (r_{t,t+kh} - d_{t,t+kh}) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_k^2).$$

If forward price data are available for n time steps back, the variance σ_k^2 can be estimated by

$$\widehat{\sigma}_k^2 = \frac{1}{n} \sum_{i=1}^{n} (s_{ih} - f_{(i-k)h,ih})^2,$$

where

$$\begin{aligned} f_{t,t+kh} &= \log(S_t \, e^{r_{t,t+kh} - d_{t,t+kh}}) \\ &= s_t + r_{t,t+kh} - d_{t,t+kh}. \end{aligned}$$

If forward price data are only available for the more recent past, the following biased estimator for σ_k^2 can be used:

$$\widehat{\sigma}_k^2 = \frac{1}{n} \sum_{i=1}^{n} (s_{ih} - s_{(i-k)h})^2.$$

If forward prices are unavailable for a desired forecast horizon kh lying between two available forecast maturities $k_1 h < k_2 h$, the trend of log-prices can be interpolated linearly

$$\mathbb{E}[s_{t+kh} - s_t \mid \mathcal{F}_t] = \frac{k_2 - k}{k_2 - k_1} \, (r_{t,t+k_1 h} - d_{t,t+k_1 h}) + \frac{k - k_1}{k_2 - k_1} \, (r_{t,t+k_2 h} - d_{t,t+k_2 h}).$$

If data are limited, constant drift and constant variance could be assumed:

$$\mathbb{E}[s_{t+kh} - s_t \mid \mathcal{F}_t] = k(r_{t,t+h} - d_{t,t+h}),$$

$$s_{t+kh} - s_t - k(r_{t,t+h} - d_{t,t+h}) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, k\sigma^2).$$

Remark: Instead of using historical volatility, option implied volatility can be used to forecast the variance. There is no explicit formula for calculating option implied volatilities, but numerical routines based on the Black-Scholes formula can be used.

3.4 Extension: value of a whole portfolio.

The above random walk based models give only distributions of individual prices and rates, but the distribution of value changes of a whole portfolio remains unknown. In this subsection we outline how a dependence structure could be added, in order to get a multivariate model.

The Vector Error Correction Model (VECM), which will be presented in Section 4, is a first approach in this direction. An alternative method is proposed in Kim et al. [26]. In the latter paper, it is suggested to construct an easier model for the distribution of the portfolio, and to apply the VECM only to estimate parameters of this easier model.

We use the following notation for the description of this method:

$$\begin{aligned}
s^l_t &= \text{log-price of asset number } l \text{ at time } t,\\
r^l_{i,j} &= s^l_{jh} - s^l_{ih},\ h\text{: time interval},\ i<j,\\
\mathbf{r}^l &= \begin{pmatrix} r^l_{0,1} \\ \vdots \\ r^l_{0,m} \end{pmatrix},\ m = \#\text{ forecast horizons},\\
\mathbf{r} &= \begin{pmatrix} \mathbf{r}^1 \\ \vdots \\ \mathbf{r}^n \end{pmatrix},\ n = \#\text{ assets},\\
k &= \overbrace{\#\text{ assets}}^{n} \cdot \overbrace{\#\text{ forecast horizons}}^{m}.
\end{aligned}$$

Extra structure is added as follows.

Dependence structure.

Assumption 1. The vector $\mathbf{r}$ of monthly asset log-returns follows a multivariate normal distribution

$$\mathbf{r} \sim \mathcal{N}_k(\widehat{\mathbf{r}}, \Sigma_{\mathbf{r}}), \tag{3.4}$$

where the vector $\widehat{\mathbf{r}}$ is non-random and depends only on forward premiums. The covariance matrix $\Sigma_{\mathbf{r}}$ can be written as

$$\Sigma_{\mathbf{r}} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & & \\ \vdots & & \ddots & \\ \Sigma_{n1} & & & \Sigma_{nn} \end{pmatrix}.$$

Note that $\Sigma^T_{ab} = \Sigma_{ba}$, but in general $\Sigma_{ab} \neq \Sigma_{ba}$.

Single price.

First we look at single prices, i.e. we are interested in the autocovariance matrices Σ_{ll} which lie on the diagonal blocks of $\Sigma_{\mathbf{r}}$.

Assumption 2. The variance depends only on the length of the time horizon:

$$\mathrm{Var}(r^l_{i,j}) = (\sigma^l_{j-i})^2, \qquad i<j. \tag{3.5}$$

With (3.5) we get for $1 \le i < j \le m$,

$$\begin{aligned}
\mathrm{Cov}(r^l_{0,i}, r^l_{0,j}) &= \mathrm{Cov}(r^l_{0,i}, r^l_{0,i}) + \mathrm{Cov}(r^l_{0,i}, r^l_{i,j})\\
&= (\sigma^l_i)^2 + \mathrm{Cov}(r^l_{0,i}, r^l_{i,j}).
\end{aligned}$$

We can express $(\sigma^l_j)^2$ in the following way,

$$\begin{aligned}
(\sigma^l_j)^2 &= \mathrm{Var}(r^l_{0,j})\\
&= \mathrm{Var}(r^l_{0,i} + r^l_{i,j})\\
&= \underbrace{\mathrm{Var}(r^l_{0,i})}_{(\sigma^l_i)^2} + \underbrace{\mathrm{Var}(r^l_{i,j})}_{(\sigma^l_{j-i})^2} + 2\,\mathrm{Cov}(r^l_{0,i}, r^l_{i,j}).
\end{aligned}$$

This means that the elements of the autocovariance matrices Σ_{ll} can be expressed by the variances:

$$\mathrm{Cov}(r^l_{0,i}, r^l_{0,j}) = \frac{1}{2}\Big((\sigma^l_i)^2 + (\sigma^l_j)^2 - (\sigma^l_{j-i})^2\Big).$$

Multiple price series.

For modelling the dependence between different price series we need the cross covariance matrices Σ_{ab}, $a \neq b$.

Assumption 3. The covariance depends only on the length of the time horizon, and there exists a correlation coefficient ρ^{ab} which is independent of time:

$$\mathrm{Cov}(r^a_{i,j}, r^b_{i,j}) = \rho^{ab}\, \sigma^a_{j-i}\, \sigma^b_{j-i}\,, \qquad i < j. \tag{3.6}$$

Correlation coefficients ρ^{ab} can be estimated by using historical log-returns r and forecasted means $\widehat{r}$ for these log-returns:

$$\widehat{\rho}^{ab} = \frac{\sum_{i=0}^{T-1}(r^a_{i,i+1} - \widehat{r}^a_{i,i+1})\,(r^b_{i,i+1} - \widehat{r}^b_{i,i+1})}{\sqrt{\sum_{i=0}^{T-1}(r^a_{i,i+1} - \widehat{r}^a_{i,i+1})^2 \sum_{i=0}^{T-1}(r^b_{i,i+1} - \widehat{r}^b_{i,i+1})^2}}\,.$$

Note that $\widehat{\rho}^{aa} = 1$, i.e. (3.6) is consistent with (3.5).

Assumption 4. The covariance depends only on the two assets a and b and on the time horizons i and j, i.e. the following symmetry holds:

$$\mathrm{Cov}(r^a_{0,j}, r^b_{0,i}) = \mathrm{Cov}(r^a_{0,i}, r^b_{0,j}) \quad \text{(which implies } \Sigma_{ab} = \Sigma_{ba}\text{)}. \tag{3.7}$$

With this assumption we get

$$\begin{aligned}
\rho^{ab}\, \sigma^a_{j-i}\, \sigma^b_{j-i} &= \mathrm{Cov}(r^a_{i,j}, r^b_{i,j}) \\
&= \mathrm{Cov}(r^a_{0,j} - r^a_{0,i}, r^b_{0,j} - r^b_{0,i}) \\
&= \mathrm{Cov}(r^a_{0,i}, r^b_{0,i}) + \mathrm{Cov}(r^a_{0,j}, r^b_{0,j}) - \mathrm{Cov}(r^a_{0,j}, r^b_{0,i}) - \mathrm{Cov}(r^a_{0,i}, r^b_{0,j}) \\
&= \underbrace{\mathrm{Cov}(r^a_{0,i}, r^b_{0,i})}_{\rho^{ab}\,\sigma^a_i\,\sigma^b_i} + \underbrace{\mathrm{Cov}(r^a_{0,j}, r^b_{0,j})}_{\rho^{ab}\,\sigma^a_j\,\sigma^b_j} - 2\,\mathrm{Cov}(r^a_{0,i}, r^b_{0,j}),
\end{aligned}$$

and finally

$$\mathrm{Cov}(r^a_{0,i}, r^b_{0,j}) = \frac{\rho^{ab}}{2}(\sigma^a_i\,\sigma^b_i + \sigma^a_j\,\sigma^b_j - \sigma^a_{j-i}\,\sigma^b_{j-i}).$$

Evaluation.

To model the multivariate price structure, we combine the distribution of individual prices with the above covariance structure. By summing up the prices of the single assets, we get the value of the whole portfolio.

4 AR(p)

In this section, we consider the AR(p) model which takes into account the dependency of subsequent returns. First we assume the following dynamics for log-prices at times

$t \in h\mathbb{N}$:

$$s_t = \sum_{i=1}^{p} a_i \, s_{t-ih} + \epsilon_t, \tag{4.1}$$

where

$$\begin{aligned} a_i &: \text{coefficients}, \\ \epsilon_t &\overset{\text{d}}{\sim} \mathcal{N}(\mu_0 + \mu_1 t, \sigma^2), \quad \epsilon_t \text{ independent}, \\ \sigma^2 &: \text{variance of the innovation process.} \end{aligned}$$

For $|\sum_{i=1}^{p} a_i| < 1$, the constant terms (like μ_0) have only a translation effect on the process s. Hence setting the constant term to zero has no effect on log-returns r. We can rewrite (4.1) as an AR(p)-process with i.i.d. $\mathcal{N}(0, \sigma^2)$ innovations:

$$\overline{s}_t = \sum_{i=1}^{p} a_i \overline{s}_{t-ih} + \overline{\epsilon}_t, \quad t \in h\mathbb{N}, \tag{4.2}$$

where

$$\begin{aligned} \overline{s}_t &= s_t - \mu t, \\ \mu &= \frac{\mu_1}{1 - \sum_{j=1}^{p} a_j}, \\ \overline{\epsilon}_t &\overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2), \\ a_j, \sigma^2 &\text{ as in (4.1).} \end{aligned}$$

4.1 Scaling rule.

For 1-period steps, the AR(p)-process (4.2) can be used to model the drift-free log-prices $\overline{s}_t = s_t - \mu t$ $(t \in h\mathbb{N})$. To calculate k-period parameters from a sample of n h-day periods, we proceed as follows:

1. Subtract the linear trend from log-prices s_t:

$$\overline{s}_t = s_t - \widehat{\mu} t \qquad \text{for } t \in h\mathbb{N},$$

   where

$$\widehat{\mu} = \frac{s_{nh} - s_0}{nh}.$$

2. Fit the AR(p)-process to the drift-free 1-period log-prices $(\overline{s}_t)_{t \in h\mathbb{N}}$: with maximum likelihood estimation (MLE) we get the estimates $\widehat{p}, \widehat{a}_i$ $(i = 1, \ldots, \widehat{p})$ and $\widehat{\sigma}_\epsilon$. Note that $\widehat{\sigma}_\epsilon$ is the volatility of the innovation process $\overline{\epsilon}$ and not of the drift-free log-prices $\overline{s}$:

$$\widehat{\sigma}_\epsilon = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (\overline{\epsilon}_{ih})^2}.$$

3. At time t, forecast the k-period log-return using the following relation:

$$\widehat{\mu}^k = kh\widehat{\mu} + \widehat{m},$$

   where

$$\widehat{m} = \widetilde{s}_{t+kh} - \widetilde{s}_t,$$

and $\widetilde{s}_{t+kh}$ is defined recursively:

$$\widetilde{s}_{t+jh} = \sum_{i=1}^{\widehat{p}} \widehat{a}_i \widetilde{s}_{t+(j-i)h} \qquad \text{for } j = 1, \dots, k,$$
$$\widetilde{s}_u = \overline{s}_u \qquad \text{for } u \le t.$$

4. Forecast the k-period volatility using the following formulas:

$$\delta_0 = 1,$$
$$\delta_j = \sum_{i=1}^{j} \widehat{a}_i \delta_{j-i}, \quad \widehat{a}_i = 0 \ \forall i > \widehat{p},$$
$$\widehat{\sigma}^k = \widehat{\sigma}_\epsilon \sqrt{\sum_{j=0}^{k-1} \delta_j^2}.$$

4.2 Estimation of value-at-risk and expected shortfall.

Like in the random walk model we can now estimate the k-period value-at-risk at a level p by

$$\widehat{\mathrm{VaR}}^p = -\big(\exp\big(\widehat{\mu}^k + \widehat{\sigma}^k x^p\big) - 1\big),$$

and the corresponding expected shortfall by

$$\widehat{\mathrm{ES}}^p = -\Big(\exp\big(\widehat{\mu}^k + \frac{(\widehat{\sigma}^k)^2}{2}\big) \frac{\Phi(x^p - \widehat{\sigma}^k)}{p} - 1\Big),$$

where

x^p : p-quantile of the standard normal distribution,

Φ : cumulative standard normal distribution function.

4.3 Extension: Vector Error Correction Model.

The Vector Error Correction Model (VECM) for modelling log-prices $\mathbf{s}_t$ of financial assets is the multi-dimensional version of an AR(p) model. The VECM is a combination of a Vector Auto Regressive Model (VARM) and an Error Correction Model (ECM). We first present these two sub models.

A VARM with p lags can be defined by the equation

$$\mathbf{s}_t = \Theta\, \mathbf{d}_t + \Gamma_1 \mathbf{s}_{t-h} + \cdots + \Gamma_p \mathbf{s}_{t-ph} + \boldsymbol{\epsilon}_t, \qquad t \in h\mathbb{N}, \tag{4.3}$$

where

$\mathbf{s}_t$: n-dimensional stochastic vector of asset log-prices,

$\mathbf{d}_t$: vector of deterministic variables, e.g. seasonal trend,

Θ : deterministic matrix of coefficients,

Γ_i : deterministic $n \times n$ matrix of coefficients,

$\boldsymbol{\epsilon}_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}_n(0, \Lambda)$,

Λ : covariance matrix.

In a VARM, each variable is explained by its own past values and the past values of all other variables.

The basic concept underlying an ECM is to make use of stable relationships between variables. Economic theory conjectures that there are long-run equilibrium forces which prevent some economic series from drifting too far apart. This long-term stability can be used to make forecasts. When there are l stable relationships in a set of n variables, we can model the independent variables with a $(n-l)$-dimensional vector ($\mathbf{v}_t = (v_t^1, v_t^2, \ldots, v_t^{n-l})^T, t \in h\mathbb{N}$) and express the original variables $\mathbf{s}_t^i$ as linear combinations of the v_t^j's plus a disturbance term. The l stable relationships provide an error correction mechanism. An ECM for n non-stationary variables and l cointegration relations can mathematically be described by

$$\begin{aligned} \mathbf{s}_t &= A\,\mathbf{v}_t + \boldsymbol{\epsilon}_t^{\mathbf{s}}, \\ \mathbf{v}_t &= \mathbf{v}_{t-h} + \boldsymbol{\epsilon}_t^{\mathbf{v}}, \end{aligned} \qquad t \in h\mathbb{N}, \tag{4.4}$$

where

$$\begin{aligned} \mathbf{s}_t, \boldsymbol{\epsilon}_t^{\mathbf{s}} &: n\text{-dimensional stochastic vectors}, \\ \mathbf{v}_t, \boldsymbol{\epsilon}_t^{\mathbf{v}} &: (n-l)\text{-dimensional stochastic vectors}, \\ A &: \text{matrix of dimension } n \times (n-l), \\ \boldsymbol{\epsilon}_t^{\mathbf{s}}, \boldsymbol{\epsilon}_t^{\mathbf{v}} &: \text{uncorrelated stationary random variables.} \end{aligned}$$

The error correction structure of this model becomes obvious, when we rewrite (4.4) in the form

$$\begin{aligned} \mathbf{s}_t &= \mathbf{s}_{t-h} - \boldsymbol{\epsilon}_{t-h}^{\mathbf{s}} + \boldsymbol{\epsilon}_t^{\mathbf{s}} + A\,(\mathbf{v}_t - \mathbf{v}_{t-h}), \\ \mathbf{v}_t &= \mathbf{v}_{t-h} + \boldsymbol{\epsilon}_t^{\mathbf{v}}, \end{aligned} \qquad t \in h\mathbb{N}. \tag{4.5}$$

The term $A\,(\mathbf{v}_t - \mathbf{v}_{t-h})$ is called the "error correction term".

Example 4.1 In order to show the idea underlying an ECM, we construct a simple example. Assume the following two-dimensional situation:

$$\begin{aligned} \mathbf{s}_t &:= \begin{pmatrix} s_t^1 \\ s_t^2 \end{pmatrix}, \\ A &:= \begin{pmatrix} a \\ 1 \end{pmatrix}, \\ \boldsymbol{\epsilon}_t^{\mathbf{s}} &:= \begin{pmatrix} \epsilon_t^{\mathbf{s},1} \\ 0 \end{pmatrix}, \\ \mathbf{v}_t &:= s_t^2, \\ \epsilon_t^{\mathbf{s},1}, \epsilon_t^{\mathbf{v}} &\overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1). \end{aligned}$$

The model (4.4) then leads to

$$\begin{aligned} s_t^1 &= a\,s_t^2 + \epsilon_t^{\mathbf{s},1}, \\ s_t^2 &= s_{t-h}^2 + \epsilon_t^{\mathbf{v}}, \end{aligned} \qquad t \in h\mathbb{N}.$$

We can rewrite this as

$$\begin{aligned} s_t^1 &= s_{t-h}^1 - \epsilon_{t-h}^{\mathbf{s},1} + \epsilon_t^{\mathbf{s},1} + a\,(s_t^2 - s_{t-h}^2), \\ s_t^2 &= s_{t-h}^2 + \epsilon_t^{\mathbf{v}}, \end{aligned} \qquad t \in h\mathbb{N}. \tag{4.6}$$

From (4.6) it is obvious that the ECM yields a possibility to model dependent stock prices. In this example, for positive values of a an increase in s_t^2 leads on average to a higher value of s_t^1. □

We can now combine the VARM and the ECM to the VECM:

$$\mathbf{r}_t = \Theta\,\mathbf{d}_t + \Pi\,\mathbf{s}_{t-h} + \sum_{i=1}^{p-1} \Gamma_i\,\mathbf{r}_{t-ih} + \boldsymbol{\epsilon}_t, \qquad t \in h\mathbb{N}, \tag{4.7}$$

where

$$\begin{aligned} \mathbf{r}_t = \mathbf{s}_t - \mathbf{s}_{t-h} &: n\text{-dimensional stochastic vector of 1-period log-returns}, \\ \mathbf{d}_t &: \text{vector of deterministic variables}, \\ \Theta &: \text{matrix of coefficients}, \\ \Gamma_i &: n \times n \text{ matrix}, \\ \Pi = \alpha\,\beta^T &: n \times n \text{ matrix}, \\ \alpha, \beta &: n \times l \text{ matrices}, \\ \boldsymbol{\epsilon}_t &\overset{\text{i.i.d.}}{\sim} \mathcal{N}_n(0, \Lambda), \\ \Lambda &: \text{covariance matrix}. \end{aligned}$$

In (4.7) the structure of the VARM is obvious. The ECM part is included in the term $\Pi\,\mathbf{s}_t$. Since the matrix Π can be decomposed into $\Pi = \alpha\,\beta^T$, we can write $\Pi\,\mathbf{s}_t$ in the form $\alpha\,\beta^T\mathbf{s}_t$, where $\beta^T\mathbf{s}_t$ are the transformed stationary series.

It remains to show that the VECM is the multi-dimensional version of an AR(p) model. In $n = 1$ dimension, the VECM (4.7) can be written as follows:

$$\begin{aligned} s_t &= s_{t-h} + \Theta\,d_t + \Pi\,s_{t-h} + \sum_{i=1}^{p-1} \Gamma_i\,(s_{t-ih} - s_{t-(i+1)h}) + \epsilon_t, \\ \epsilon_t &\overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2). \end{aligned} \tag{4.8}$$

By substituting $a_1 := 1 + \Pi + \Gamma_1$, $a_i := \Gamma_i - \Gamma_{i-1}$ $(i = 2, 3, \ldots, p-1)$, $a_p := -\Gamma_{p-1}$, and by assuming (d_t) to be linear in time (i.e. $\Theta\,d_t = \mu_0 + \mu_1 t$), we get (4.1). This model we rewrote in (4.2) as an AR(p)-process with i.i.d. $\mathcal{N}(0, \sigma^2)$ innovations.

5 GARCH(1,1)

Let $(\overline{r}_t,\, t \in h\mathbb{N})$ be a strictly stationary time series representing observations of the centered 1-period log-returns of a financial asset. We point out that we do not specify yet at which frequency we observe the returns: it can vary from daily to yearly frequencies.

We assume the dynamics of $\overline{r}$ to be given by a GARCH(1,1) process with Student-t distributed innovations:

$$\begin{aligned} \overline{r}_t &= \sigma_t \, \epsilon_t \quad \text{for } t \in h\mathbb{N}, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 \overline{r}_{t-h}^2 + \beta_1 \sigma_{t-h}^2, \\ \epsilon_t &\overset{\text{i.i.d.}}{\sim} t_\nu, \, \mathbb{E}[\epsilon_t] = 0, \, \mathbb{E}[\epsilon_t^2] = 1, \end{aligned} \tag{5.1}$$

where σ_t is assumed to be $\mathcal{F}_{t-h}$-measurable, and ϵ_t is the innovation process. We assume that the stationarity conditions, $0 < \alpha_0 < \infty$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$ and $\alpha_1 + \beta_1 < 1$, are fulfilled.

This model is known to be well-suited for reproducing the heteroscedastic behaviour of the conditional volatility of financial daily returns. In practice, it is intensively used for estimating the 10-day value-at-risk. One often fits the model to daily data, and then applies the square-root-of-time rule to extend the results to a longer time horizon (e.g. 10 days). For a 10-day horizon, Brummelhuis and Kaufmann [7] show that the square-root-of-time rule works well to scale a 1-day value-at-risk to a 10-day value-at-risk in a GARCH(1,1) model. However, this scaling is inappropriate when applied to long term horizon conversion. Christoffersen et al. [10] discuss the square-root-of-time scaling rule, showing that in GARCH(1,1) models for large k (such as $k > 15$ days) applying this standard rule to the strongly varying short-time volatilities σ_t leads to an overestimation of (actually low) volatility fluctuations of k-period log-returns (see also McNeil and Frey [29] for a critical discussion on this scaling rule).

We suggest to consider lower frequency data, as fortnightly or monthly returns to estimate the parameters of the GARCH(1,1) model, and then to use the temporal aggregation rule described in detail in the following section to get the one-year risk estimate. We resort to the pseudo-maximum-likelihood estimation to get the value of the parameters of the conditional volatility in (5.1).

5.1 Scaling rule.

Drost and Nijman [13] define and investigate the temporal aggregation of so-called weak GARCH processes. We refer to their paper for the formal definition of this version of GARCH processes. The so-called Drost-Nijman formula allows to extend short-term risk estimates to a one-year horizon. We suppose that the centered 1-period log-return series $(\overline{r}_t, t \in h\mathbb{N})$ are modelled by a GARCH(1,1) process as defined in equation (5.1), with symmetric innovations. Drost and Nijman [13] show that, under regularity conditions, the corresponding centered k-period log-returns $(\overline{r}_t^k, \, t \in kh\mathbb{N})$ follow a weak GARCH(1,1) process with

$$(\sigma_t^k)^2 = \alpha_{k,0} + \alpha_{k,1} \, (\overline{r}_{t-kh}^k)^2 + \beta_{k,1} \, (\sigma_{t-kh}^k)^2, \tag{5.2}$$

where

$$\begin{aligned} \alpha_{k,0} &= k \, \alpha_0 \frac{1 - (\alpha_1 + \beta_1)^k}{1 - (\alpha_1 + \beta_1)}, \\ \alpha_{k,1} &= (\alpha_1 + \beta_1)^k - \beta_{k,1}, \end{aligned}$$

and $|\beta_{k,1}| < 1$ is the solution of the quadratic equation

$$\frac{\beta_{k,1}}{1+(\beta_{k,1})^2} = \frac{a(\alpha_1+\beta_1)^k - b}{a(1+(\alpha_1+\beta_1)^{2k}) - 2b}\,.$$

Further,

$$a = k\,(1-\beta_1)^2 + 2k\,(k-1)\,\frac{(1-\beta_1-\alpha_1)^2(1-\beta_1^2-2\alpha_1\beta_1)}{(\widetilde{\kappa}-1)(1-(\alpha_1+\beta_1)^2)} + 4\,\frac{(k-1-k(\alpha_1+\beta_1)+(\alpha_1+\beta_1)^k)(\alpha_1-\alpha_1\beta_1(\alpha_1+\beta_1))}{1-(\alpha_1+\beta_1)^2}\,,$$

$$b = (\alpha_1 - \alpha_1\,\beta_1\,(\alpha_1+\beta_1))\,\frac{1-(\alpha_1+\beta_1)^{2k}}{1-(\alpha_1+\beta_1)^2}\,,$$

and $\widetilde{\kappa}$ is the unconditional kurtosis of $\overline{r}_t$.

Since $\alpha_1+\beta_1 < 1$ (and $\alpha_1 \geq 0$, $\beta_1 \geq 0$), the stationarity condition $0 \leq \alpha_{k,1}+\beta_{k,1} < 1$ is fulfilled, and we get the following asymptotic behaviour:

$$\lim_{k\to\infty} \frac{\alpha_{k,0}}{k} = \frac{\alpha_0}{1-(\alpha_1+\beta_1)}\,.$$

From the formulas for $\alpha_{k,1}$ and $\beta_{k,1}$ we can see that

$$\alpha_{k,1} \to 0, \beta_{k,1} \to 0 \ \text{ as } k \to \infty. \tag{5.3}$$

Therefore, long term volatility fluctuations stay rather small. More precisely, asymptotically the volatility becomes constant, and therefore the weak GARCH(1,1) process behaves in the limit like a random walk.

The unconditional kurtosis $\widetilde{\kappa}^k$ of $\overline{r}_t^k$ is given by

$$\widetilde{\kappa}^k = 3 + \frac{\widetilde{\kappa}-3}{k} + 6(\widetilde{\kappa}-1)\frac{(k-1-k(\alpha_1+\beta_1)+(\alpha_1+\beta_1)^k)(\alpha_1-\alpha_1\beta_1(\alpha_1+\beta_1))}{k^2(1-\alpha_1-\beta_1)^2(1-\beta_1^2-2\alpha_1\beta_1)}. \tag{5.4}$$

Note that formula (5.4) describes how unconditional kurtosis changes when the time horizon is enlarged by a factor k. This equation cannot be applied directly to the *conditional* normalized fourth moments. But we can use the following relations between unconditional kurtosis $\widetilde{\kappa}$ (respectively $\widetilde{\kappa}^k$ for k periods) and the conditional kurtosis κ (respectively κ^k for k periods):

$$\widetilde{\kappa} = \frac{1-(\alpha_1+\beta_1)^2}{1-(\alpha_1+\beta_1)^2-{\alpha_1}^2\,(\kappa-1)}\,\kappa, \tag{5.5}$$

and

$$\kappa^k = \frac{1-(\alpha_1+\beta_1)^2+\alpha_1^2}{1-(\alpha_1+\beta_1)^2+\alpha_1^2\,\widetilde{\kappa}^k}\,\widetilde{\kappa}^k. \tag{5.6}$$

Proofs of the two equations (5.5) and (5.6) can be found in Appendix A.1. Note that the Drost-Nijman scaling rule does not give exact information about the distribution of the innovations ϵ_t^k. Only even moments are calculated, indeed, all odd moments are zero because of the assumption of symmetric one period innovations. We can model the innovations ϵ_t^k with Student-t distributions with correct first five moments (i.e. two even moments). When we assume Student-t distributions for one-period innovations ϵ_t, the estimation procedure for the k-period parameter ν^k consists of the following steps:

1. estimate the parameter ν of the Student-t distributed one-period innovations,
2. calculate the conditional kurtosis κ via $\kappa = \frac{3\nu-6}{\nu-4}$,
3. calculate the unconditional kurtosis $\widetilde{\kappa}$ via (5.6),
4. apply (5.4) to get the unconditional kurtosis $\widetilde{\kappa}^k$ for k-period log-returns,
5. calculate the conditional kurtosis κ^k for k-period log-returns via (5.5),
6. calculate the parameter ν^k of the Student-t distributed innovations (ϵ_t^k) via $\nu^k = \frac{4\kappa^k-6}{\kappa^k-3}$.

5.2 Estimation of value-at-risk and expected shortfall.

We would like to work with a GARCH(1,1) model with a Student-t distributed innovation process for modelling centered 1-period log-returns $\overline{r}_t$, i.e. to use $\epsilon_t \overset{\text{i.i.d.}}{\sim} t_\nu$ in model (5.1). We fit the GARCH(1,1) process to 1-period (h-day) log-returns using quasi maximum likelihood estimators (QMLE). We use the S-Plus package GARCH to carry out these estimations. Unfortunately, this S-Plus estimation routine is not stable when working with Student-t distributed innovations. Therefore, for the estimation procedure, we simplify our model and continue working with standard normally distributed innovations for 1-period log-returns:

$$\begin{aligned}
\overline{r}_t &= \sigma_t \, \epsilon_t, \\
\sigma_t^2 &= \alpha_0 + \alpha_1 \overline{r}_{t-h}^2 + \beta_1 \sigma_{t-h}^2, \\
\epsilon_t &\overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1).
\end{aligned}$$

After estimating $\widehat{\alpha}_0$, $\widehat{\alpha}_1$ and $\widehat{\beta}_1$, we apply the Drost-Nijman scaling rule to calculate $\widehat{\alpha}_{k,0}$, $\widehat{\alpha}_{k,1}$ and $\widehat{\beta}_{k,1}$ for the centered k-period log-returns. For estimating ν^k, we use the six steps listed above. In step 1, we take the limit $\nu \to \infty$, which corresponds to assuming standard normally distributed 1-period innovations ϵ, and which yields $\kappa = 3$ in step 2.

An alternative approach to taking $\nu \to \infty$ would be to work with Student-t distributed 1-period innovations, and to only assume normality for the estimation procedure (to get $\widehat{\alpha}_0$, $\widehat{\alpha}_1$ and $\widehat{\beta}_1$). The parameter ν can then be estimated from the residuals. Gouriéroux [23] shows that this so-called pseudo-maximum-likelihood method yields a consistent and asymptotically normal estimator.

At time t we forecast the k-period conditional volatility $\widehat{\sigma}^k = \widehat{\sigma}(t,t)$ using the following recursive relation: for $t^* = t-(n-1)kh, \dots, t-kh, t \quad (n = \#\ k\text{-period log-returns})$

$$\widehat{\sigma}^2(t^*,t) = \widehat{\alpha}_0 + \widehat{\alpha}_1 (r_{t^*}^k - \widehat{\mu}^k)^2 + \widehat{\beta}_1 \widehat{\sigma}^2(t^*-kh,t),$$

starting with

$$\widehat{\sigma}^2(t-nkh,t) = k \, \frac{1}{nk-1} \sum_{i=0}^{nk-1} (r_{t-ih} - \widehat{\mu})^2.$$

The k-period mean can be estimated by

$$\widehat{\mu}^k = k\widehat{\mu},$$

where $\widehat{\mu}$ is the QMLE for the mean 1-period log-return up to time t.

Since we assumed Student-t distributed k-period innovations, we can estimate the k-period (one-year) value-at-risk at a level p by

$$\widehat{\text{VaR}}^p = -\big(\exp\big(\widehat{\mu}^k + \widehat{\sigma}^k\, x^p_{\widehat{\nu}^k}\big) - 1\big),$$

and the corresponding expected shortfall by

$$\widehat{\text{ES}}^p = -\left(\frac{1}{p}\int_0^p \exp(\widehat{\mu}^k + \widehat{\sigma}^k\, x^q_{\widehat{\nu}^k})dq - 1\right),$$

where x^p_ν is the p-quantile of a Student-t distributed random variable with parameter ν, mean zero and variance one. The integral can be evaluated numerically.

5.3 Extension: models of changing volatility.

The general structure of models we discuss in this section is as follows:

$$\overline{r}_t = \sigma_t\, \epsilon_t, \quad \mathbb{E}[\epsilon_t] = 0,\, \mathbb{E}[\epsilon_t^2] = 1, \quad t \in h\mathbb{N}, \tag{5.7}$$

where σ_t is a stochastic process. Special cases are (see Campbell et al. [8] for more details):

- Markovian stochastic volatility models: Todays volatility only depends on the volatility one period ago. Hence

$$\sigma_t = F(\sigma_{t-h}, Z_t, t), \quad t \in h\mathbb{N},$$
$$Z_t = \text{white noise (i.e. i.i.d.), independent of } (\epsilon_t)_{t\in h\mathbb{N}}.$$

- Generalisation of Markovian stochastic volatility models: There is additionally dependence on the return one period ago:

$$\sigma_t = F(\sigma_{t-h}, \overline{r}_{t-h}, Z_t, t), \quad t \in h\mathbb{N}.$$

- In regime-switching volatility models, volatility is state dependent:

$$\sigma_t = v_{U_t}, \quad U_t\text{: state at time } t,$$

  where the transition probabilities are given by a matrix Π with

$$\Pi_{u_1,u_2} = \mathbb{P}[U_{t+1} = u_2 \mid U_t = u_1].$$

- Log-auto-regressive volatility:

$$\log \sigma_t^2 = \alpha + \gamma \log \sigma_{t-h}^2 + \kappa\, Z_t, \quad t \in h\mathbb{N},$$

$$Z_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1).$$

- GARCH(p,q)

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i\, \overline{r}_{t-ih}^2 + \sum_{j=1}^q \beta_j\, \sigma_{t-jh}^2, \quad t \in h\mathbb{N}.$$

  Details on the GARCH(1,1) model were given before.

- EGARCH(1,1)

$$\log \sigma_t^2 = \alpha + \gamma \, \log \sigma_{t-h}^2 + \beta_1 \frac{\overline{r}_{t-h}}{\sigma_{t-h}} + \beta_2 \left(\left| \frac{\overline{r}_{t-h}}{\sigma_{t-h}} \right| - \sqrt{\frac{2}{\pi}} \right), \quad t \in h\mathbb{N},$$

  and their generalisation for general p and q.

- Cross-market GARCH: adding a dependence structure between volatilities in different markets.

  A 2-market-version takes the following form:

$$\begin{pmatrix} (\sigma_t^a)^2 \\ \sigma_t^{ab} \\ (\sigma_t^b)^2 \end{pmatrix} = \alpha + \beta \begin{pmatrix} (\overline{r}_{t-h}^a)^2 \\ \overline{r}_{t-h}^a \, \overline{r}_{t-h}^b \\ (\overline{r}_{t-h}^b)^2 \end{pmatrix} + \gamma \begin{pmatrix} (\sigma_{t-h}^a)^2 \\ \sigma_{t-h}^{ab} \\ (\sigma_{t-h}^b)^2 \end{pmatrix}, \quad t \in h\mathbb{N},$$

$$\begin{aligned} \overline{r}_t^a &: \text{centered log-return in market } a \text{ at time } t, \\ \overline{r}_t^b &: \text{centered log-return in market } b \text{ at time } t, \\ \sigma_t^a &: \text{conditional volatility of } \overline{r}_t^a, \\ \sigma_t^b &: \text{conditional volatility of } \overline{r}_t^b, \\ \sigma_t^{ab} &: \text{conditional covariance between } \overline{r}_t^a \text{ and } \overline{r}_t^b, \\ \beta, \gamma &: \text{matrices of dimension } 3 \times 3. \end{aligned}$$

- As a generalisation of (5.7), a model of the following form can be used:

$$\overline{r}_t = \sigma_t \, \epsilon_t + N_t \, \sigma_t^* \, \epsilon_t^*, \quad t \in h\mathbb{N},$$

  where

$$\begin{aligned} \sigma_t, \sigma_t^* &: \text{ deterministic or stochastic,} \\ \epsilon_t & \text{ i.i.d., } \ \mathbb{E}[\epsilon_t] = 0, \, \mathbb{E}[\epsilon_t^2] = 1, \\ \epsilon_t^* & \text{ i.i.d., independent of } (\epsilon_t)_{t \in h\mathbb{N}}, \\ N_t & \overset{\text{i.i.d.}}{\sim} \text{Be}(p). \end{aligned}$$

  Be(p) denotes a Bernoulli distribution with parameter p. Defined like this, the arrival process of N_t is a Poisson process. The distribution of ϵ_t^* is usually skewed. This generalisation allows to model short-term, sudden changes in the volatility process (similar to jumps in a continuous-time model).

6 Heavy-tailed distributions

Extreme Value Theory (EVT) has attracted much interest from the finance industry lately. It typically offers a semi-parametric estimation procedure for extreme tail estimation. For an exhaustive description of EVT see the monograph of Embrechts et al. [19]. Applications to risk management are summarised in Embrechts [18].

We consider $(r_t, t \in h\mathbb{N})$ i.i.d., representing observations of the log-returns. Further, we assume the following asymptotic behaviour for the left tail:

$$\mathbb{P}[r < -x] = x^{-\alpha} L(x) \quad \text{as } x \to \infty, \tag{6.1}$$

where $\alpha \in \mathbb{R}^+$ and L is a slowly varying function, i.e. $\forall t > 0$

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1.$$

Distributions fulfilling (6.1) are often called "heavy-tailed distributions". The mth moment is infinite for $m > \alpha$. A natural estimator for the tail index α is the so-called Hill estimator

$$\widehat{\alpha}_{l,n}^{-1} = \frac{1}{l} \sum_{i=1}^{l} \log \left(\frac{r_{(i)}}{r_{(l)}} \right), \tag{6.2}$$

where n is the sample size and $r_{(l)}$ is the lth order statistic (i.e. $r_{(1)} \leq r_{(2)} \leq \cdots \leq r_{(n)}$), taken as a threshold.

From (6.1) we can derive the following estimates for the p-quantile $\widehat{x}_{l,n}^p$:

- Putting $x = -r_{(l)}$ we get

$$\begin{aligned} \frac{l}{n} &= \mathbb{P}[r < r_{(l)}] \\ &= (-r_{(l)})^{-\alpha} L(-r_{(l)}), \end{aligned}$$

- and setting $x = -x_{l,n}^p$ yields

$$\begin{aligned} p &= \mathbb{P}[r < x_{l,n}^p] \\ &= (-x_{l,n}^p)^{-\alpha} L(-x_{l,n}^p). \end{aligned} \tag{6.3}$$

Combining these equations leads to the estimate

$$\widehat{x}_{l,n}^p = r_{(l)} \left(\frac{l}{np} \right)^{1/\widehat{\alpha}_{l,n}}. \tag{6.4}$$

Formula (6.4) yields an easy way to calculate the p-quantile once the tail (shape) parameter α is estimated. As the Hill estimator (6.2) is based upon the highest (lowest) order statistics, we consider only a proportionally small subsample of the whole data set. Consequently, a large data set to start with is desirable. Therefore, we use quite high frequency data to estimate α and apply a theoretical scaling rule to get the quantile estimate at the required frequency.

6.1 Scaling rule.

We still assume 1-period log-returns to be i.i.d., satisfying (6.1). By the subexponentiality property (see e.g. Feller [22, VIII.8, p.271]), we get for the k-period log-returns:

$$\mathbb{P}\left[r^k < -x\right] = k\, x^{-\alpha} L(x)(1 + o(1)), \quad \text{as } x \to \infty. \tag{6.5}$$

When applicable, this result supplements the central limit theorem by providing information concerning the tails. It describes the self-additivity in the tails of heavy-tailed distributions. The implication of this result for portfolio analysis has been discussed in the specific case of non-normal stable distributions by Fama and Miller [21]. In that case $\alpha < 2$ holds, and the variance is therefore infinite. Here we focus on the finite variance case. Dacorogna et al. [11] show the following asymptotic result, which can be derived from (6.3) and (6.5):

$$x^{k,p} \sim k^{1/\alpha} x^p, \quad \text{as } p \to 0, \tag{6.6}$$

where $x^{k,p}$ is the p-quantile of the k-period log-returns. This means that at a constant probability p, increasing the time horizon by a factor k increases the p-quantile for the heavy-tailed model by a factor $k^{1/\alpha}$. For financial log-returns, the estimated tail index α usually belongs to the interval $(2,5)$ (see Dacorogna et al. [11] and Straetmans [35] for empirical studies on foreign exchange rate and stock index data). For $\alpha > 2$ the factor $k^{1/\alpha}$ is smaller than for the normal model, where the value-at-risk is increased by the factor $k^{1/2}$. Hence the square-root-of-time rule – used very often in practice to scale quantiles – leads to an overestimation of value-at-risk.

In comparison with the normal model, the probability of an extreme 1-period loss is higher for the heavy tailed model. However, as we saw before, the multiplication factor used to obtain the multi-period value-at-risk is smaller for heavy-tailed log-returns ($k^{1/\alpha}$, $\alpha \in (2,5)$) than for normal log-returns ($k^{1/2}$). Reasoned by these two counterbalancing forces, the value of an estimated extreme k-period quantile might be larger for the normal model than for the EVT model, if k is chosen large enough (as shown by simulation in Dacorogna et al. [11]).

6.2 Estimation of value-at-risk and expected shortfall.

We assume that the 1-period log-returns are independent, and that

$$\mathbb{P}[r < -x] = x^{-\alpha} L(x) \quad \text{as } x \to \infty, \tag{6.7}$$

where $\alpha \in \mathbb{R}^+$ and L is a slowly varying function.

We apply the quantile transformation (6.6) to 1-period values for deriving estimates for the k-period value-at-risk at a level p

$$\widehat{\mathrm{VaR}}^p = -\Big(\exp\Big(\big(\frac{k\, l(n,p)}{n\,p}\big)^{1/\widehat{\alpha}_{l(n,p),n}}\; r_{(l(n,p))}\Big) - 1\Big),$$

and for the corresponding one-year expected shortfall

$$\widehat{\mathrm{ES}}^p = -\Big(\frac{1}{p}\int_0^p \exp\Big(\big(\frac{k\, l(n,p)}{n\,q}\big)^{1/\widehat{\alpha}_{l(n,p),n}}\; r_{(l(n,p))}\Big)\, dq - 1\Big),$$

where

$$l(n,p) = \lfloor n(p + 0.045 + 0.005\, h)\rfloor,$$

$$\widehat{\alpha}_{l,n}^{-1} = \frac{1}{l}\sum_{i=1}^{l} \log\left(\frac{r_{(i)}}{r_{(l)}}\right) \qquad \text{(Hill estimator)},$$

$r_{(l)}$: the lth order statistic of 1-period log-returns, i.e. $r_{(1)} \le r_{(2)} \le \cdots \le r_{(n)}$.

The integral in the estimation of $\widehat{\mathrm{ES}}^p$ can be evaluated numerically. The choice made for $l(n,p)$ works well in practice; it turns out that the results are rather insensitive on the exact choice of $l(n,p)$.

6.3 A specific case: the generalized inverse Gaussian distributions.

In this section we describe the generalized inverse Gaussian distribution which belongs to the class of generalized hyperbolic distributions. This class of distributions has been introduced by Barndorff-Nielsen [4] and studied for financial applications by Eberlein and Keller [16]. The two most attractive features of a generalized inverse Gaussian distribution for use in long term risk management are on the one hand its flexibility due

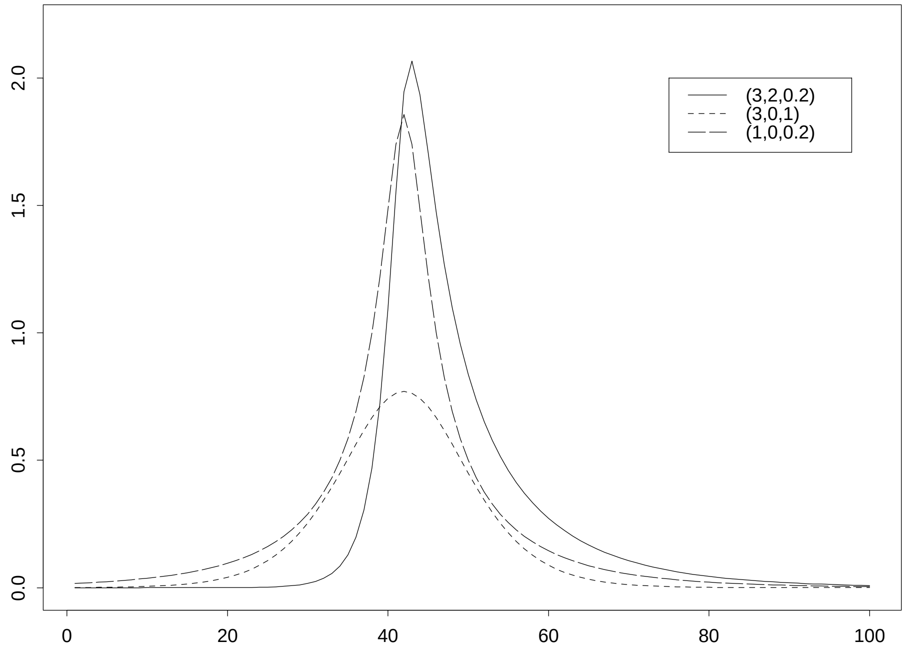

Figure 1 Densities of the Normal Inverse Gaussian distribution for different values of the parameters (α, β, δ).

to the five parameters driving the model, and on the other hand the stability property under convolution. Indeed, the flexibility allows to fit financial data well at any frequency. The time aggregate distribution for longer time horizons can be obtained by convolution.

The density of a generalized inverse Gaussian distribution is given by

$$f_{GH}(x; \lambda, \alpha, \beta, \delta, \mu) \\ = a(\lambda, \alpha, \delta, \mu)(\delta^2 + (x-\mu)^2)^{(\lambda-\frac{1}{2})/2} K_{\lambda-\frac{1}{2}}\Big(\alpha\sqrt{\delta^2 + (x-\mu)^2}\exp(\beta(x-\mu))\Big), \tag{6.8}$$

where

$$a(\lambda, \alpha, \beta, \delta) = \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^\lambda K_\lambda(\delta\sqrt{\alpha^2-\beta^2})}$$

is a normalizing constant and K_ν denotes the modified Bessel function of the third kind with index ν. An integral representation of K_ν is given by

$$K_\nu(z) = \frac{1}{2}\int_0^\infty y^{\nu-1}\exp\Big(-\frac{1}{2}z(y+y^{-1})\Big)dy.$$

The densities above depend on five parameters:

- $\alpha > 0$ determines the shape,
- β with $0 \le |\beta| < \alpha$ the skewness,
- $\mu \in \mathbb{R}$ the location,
- $\delta > 0$ is a scaling parameter, and
- $\lambda \in \mathbb{R}$ characterizes certain sub-classes (the tail behaviour can be modified by changing λ).

This class of distributions is invariant under affine transformations. The special case $\lambda = -\frac{1}{2}$ corresponds to the normal inverse Gaussian distribution:

$$f_{NIG}(x) = \frac{\alpha}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} + \beta(x-\mu)\right) \frac{K_1\left(\alpha\delta\sqrt{1+\left(\frac{x-\mu}{\delta}\right)^2}\right)}{\sqrt{1+\left(\frac{x-\mu}{\delta}\right)^2}}. \qquad (6.9)$$

In order to show the flexibility of the NIG distributions, we have plotted its density for different values of the parameters (see Figure 1). The class of normal inverse Gaussian distributions deserves a particular interest because it is the only sub-class which is stable under convolution. It fulfils the following aggregation property:

$$NIG(\alpha,\beta,\delta_1,\mu_1) * NIG(\alpha,\beta,\delta_2,\mu_2) = NIG(\alpha,\beta,\delta_1+\delta_2,\mu_1+\mu_2). \qquad (6.10)$$

In this sense it is close to the normal distribution where means and variances of independent random variables add up as well. Eberlein [15] gives analytical properties and in particular the moment generating function which allows to derive different moments. A parameter estimation procedure is described in Prause [33].

7 Empirical studies

In this section, we first investigate the statistical properties and the behaviour of financial data considered at different frequencies (from 1 day to 30 days). Then we carry out some tests concerning the temporal aggregation of returns. These two studies are essential to understand better the performance of each of the models described in the previous parts. As for an institutional investor the data sets of main interest are exchange rates, stock indices and government bonds, we look at data in each of these categories.

7.1 Data exploration.

7.1.1 *Figures.*

Table 1 (foreign exchange rates from 01.01.1985 to 29.12.2000), Table 2 (stock indices from 01.01.1990 to 29.12.2000) and Table 3 (10-year government bonds from 01.01.1990 to 29.12.2000) present some descriptive statistics of log-returns taken at different frequencies from one day to one month. We remark the well-known stylized facts of financial log-returns: leptokurtosis and skewness. The other typical observations which should be mentioned are in particular the decrease of kurtosis with respect to frequency for most data sets, whereas there is evidence for persistence of skewness.

From now on, we explore two data sets more intensively: the USD/DEM exchange rate from 03.01.1980 to 21.05.1996 and the German Stock Index (Deutscher Aktienindex, DAX) from 02.01.1973 to 23.07.1996, see Figure 3. We carry on our data exploration by using two graphical tools, the quantile-quantile-plot and the sample autocorrelation function.

7.1.2 *The quantile-quantile-plot (QQ-plot) against the normal distribution.*

The QQ-plot is defined by

$$\left\{ \left(r_{(l)}, \Phi^{-1}\left(\frac{n-l+1}{n+1} \right) \right),\ l = 1, \ldots, n \right\},$$

where $r_{(l)}$ denotes the lth order statistic and Φ^{-1} is the inverse of the cumulative standard Gaussian distribution function. The QQ-plots in Figures 4 and 5 confirm that daily data are clearly heavier tailed than normally distributed data, whereas monthly data seem to be almost as light tailed as normally distributed data. As opposed to the USD/DEM data, the DAX sample is clearly skewed at all evaluated frequencies.

7.1.3 *The sample autocorrelation function: long range dependence or non-stationarity?*

Though financial log-returns do not exhibit clear correlation, this is not the case for the absolute and square value of log-returns. We focus our investigation only on the sample autocorrelations of absolute values. In this case we do not have to assume finite fourth moments of financial data. Finite second moments are sufficient. Indeed some log-returns of financial data do have infinite third or fourth moment (see Embrechts et al. [19] for the statistical theory of tail and high quantile estimation).

Assuming $\mathbb{E}[|r_t^2|] < \infty$, we plot

$$\left\{ \left(h, \widetilde{\rho}_{n,|\mathbf{r}|}(k) \right),\ k = 1, 2, \ldots, 100 \right\},$$

where

$$\widetilde{\rho}_{n,|\mathbf{r}|}(k) = \frac{\gamma_{n,|\mathbf{r}|}(k)}{\sqrt{v_{\mathbf{r}}(k)}}, \qquad k \in \mathbb{N},$$

and

$$\gamma_{n,|\mathbf{r}|}(k) = \left(\frac{1}{n} \sum_{t=1}^{n-k} |r_t||r_{t+k}| \right) - \left(\mathbb{E}[|\mathbf{r}|] \right)^2, \qquad k \in \mathbb{N},$$

$$v_{\mathbf{r}}(k) = \mathbb{E}[(r_0 r_k)^2], \qquad k \in \mathbb{N}.$$

We plot the sample autocorrelation functions for different periods. In the Figures 6 and 7, we observe that for samples of long periods the sample autocorrelations are significantly different from zero with positive values, even for large lags. On the one hand, this empirical finding is usually interpreted as evidence for long memory in the volatility of financial log-returns. On the other hand, Mikosch and Stărică [30] as well as Diebold and Inoue [12] show theoretically that similar effects can be observed for time series which consist of subsamples originating from different stationary models, so-called regime switching models. Mikosch and Stărică conclude: "It might be misleading to take the empirical evidence of long memory and strong persistence of the volatility in log-returns at face value, especially when it comes from the analysis of time series that cover long periods". Indeed, we also observe changes in the sample autocorrelations when we consider time windows of different sizes. It can be seen in Figure 7 that for short periods (lower graphs) values are small from a lag of 10 days onward. An explanation of this observation might be a type of non-stationarity, i.e. shifts in the unconditional volatility of the model underlying the log-returns (c.f. regime switching models in Section 5.3).

7.2 Scaling in practice.

For the approaches described in Sections 3 to 6, we fit our models to higher frequency data (e.g. one day to one month) than required for the long-term risk horizon under study (e.g. one year). This makes sense, since at some point we have to make a bias-variance tradeoff. Indeed, considering too high frequency data would lead to a bias in the risk measures (see Hancock et al. [24]), whereas using too low frequency data will lead to high variability for the parameters fitted, caused by the lack of data.

7.2.1 *Empirical scaling function.*

In order to verify theoretical time rules for temporal risk aggregation (like the square-root-of-time rule, the EVT scaling rule, etc.), and also to make sure that empirical results are reasonable, it is necessary to explore financial data. Before we start comparing the models in Section 8, we investigate the empirical scaling of quantiles. We proceed as follows. Let $(r_t)_{1\leq t\leq n}$ represents n observations of daily log-returns on a financial asset. We denote by $x^{k,p}$ the quantile defined by

$$x^{k,p} = \inf\{r \in \mathbb{R} \mid \mathbb{P}[\sum_{i=1}^{k} r_i \leq r] \geq p\}. \tag{7.1}$$

We are interested in finding an empirical relation between quantile values computed at the same confidence level p, for different frequencies. Put differently, we look for a relation between different convolutions of our time series. We expect this relation to be dependent on the convolution order k. Moreover, we assume that this function with respect to time is of power type, like for i.i.d. normally distributed or heavy tailed random variables. Formally, we can write this relation as follows:

$$x^{k,p} = \left(\frac{k}{k^*}\right)^{\beta_{k,k^*}} x^{k^*,p}, \qquad 1 \leq k^* \leq k,\ \beta_{k,k^*} \in \mathbb{R},\ \text{and } p \in [0,1]. \tag{7.2}$$

Thus,

$$\beta_{k,k^*} = \log\left(\frac{x^{k,p}}{x^{k^*,p}}\right)\left(\log\left(\frac{k}{k^*}\right)\right)^{-1}, \qquad 1 \leq k^* \leq k,\ \text{and } p \in [0,1]. \tag{7.3}$$

As reference quantile $x^{k^*,p}$ we choose the quantile obtained from the empirical distribution of daily log-returns (i.e. $k^* = 1$). Then, from our data set of size n we start by picking the first $m \leq n$ data, we estimate quantiles for different values of k ($k = 2, \ldots, y$), and finally we compute β_{k,k^*} by using (7.3). We shift our time window and repeat this procedure $l = n-m+1$ times (overlapping technique) to obtain values $\beta^{(i)}_{k,k^*}$ ($i = 1, \ldots, l$). For fixed k we denote their average by

$$\bar{\beta}_{k,1} = \frac{1}{l}\sum_{i=1}^{l} \beta^{(i)}_{k,1}. \tag{7.4}$$

We plot for each data set (with $p = 0.05$):

$$\left\{(k, \bar{\beta}_{k,1}),\ k = 2, \ldots, y\right\}.$$

For the Figures 8 and 9, we have used non overlapping data for estimating the quantiles. We observe that the scaling factor β of quantiles of index data has a tendency to decrease (from 0.7 to 0.4) with the time horizon which runs from 1 to 26 days. For the FX data, there is no clear trend since β fluctuates around 0.5.

When we compare Figure 8 with the data analysis in Table 2, we notice that there are similarities between the kurtosis and the power parameter behaviour with respect to time: as the time horizon increases, the kurtosis and the power parameter decrease simultaneously.

7.2.2 *Test of the Drost-Nijman formula.*

In this part, we examine whether the Drost-Nijman formula (5.2) gives good approximations for real data. We use quasi maximum likelihood estimators (QMLE) to get the GARCH(1,1) parameters $\alpha_0, \alpha_1, \beta_1$ (assuming normally distributed innovations) for centered daily log-returns. For the DAX this leads to

$$\sigma_t^2 = 2.750 \cdot 10^{-6} + 0.09706\, \overline{r}_{t-1}^2 + 0.8815\, \sigma_{t-1}^2,$$

and for the USD/DEM exchange rate we get

$$\sigma_t^2 = 4.472 \cdot 10^{-7} + 0.05127\, \overline{r}_{t-1}^2 + 0.9393\, \sigma_{t-1}^2.$$

We have now two possibilities for estimating the GARCH(1,1) parameters of the centered h-day log-returns. We can either use the temporal aggregation formula (5.2) of GARCH(1,1) processes for a good approximation of quantiles, or we can estimate the parameters directly with data taken at the requested frequency. In the latter case we once again assume the h-day innovations to be normally distributed.

From (5.1) we can see that in a GARCH(1,1) model for daily log-returns the conditional variance can be rewritten as $\mathbb{E}[\sigma_t^2 \mid \sigma_{t-1}^2] = \alpha_0 + (\alpha_1 + \beta_1)\, \sigma_{t-1}^2$. Therefore the most important parameter is the sum $(\alpha_1 + \beta_1)$. In Tables 4 and 5 we observe that for a time horizon of five days the estimated parameters are quite close to the ones calculated via Drost-Nijman. For the USD/DEM example, this remains true for a horizon of 20 days. For longer horizons the estimates differ significantly. A main reason for this is the fact that the data sets are too short to enable reliable estimates via the QMLE method. Also the stationarity assumption for such long periods is quite unrealistic.

In order to trace the origin of the problems for GARCH(1,1) models, we repeat this test for a simulated time series. We assume the following volatility process:

$$\sigma_t^2 = 2 \cdot 10^{-6} + 0.08\, \overline{r}_{t-1}^2 + 0.90\, \sigma_{t-1}^2.$$

From this, we simulate 5000 daily returns (which is between the number of returns we have for the DAX, and for the USD/DEM exchange rate, respectively). We apply again the two methods (scaling formula and direct estimation) to get GARCH(1,1) parameters for longer horizons, see Table 6. These results are in line with the ones for the DAX and for the USD/DEM exchange rate. Hence it is not only non-stationarity (which is not present in these simulated data) which causes problems. The lack of sufficiently many h-day returns turns out to be at least as unfavourable.

As we just saw, for short horizons the results from direct estimation indicate appropriateness of the Drost-Nijman formula. Due to the problems when working with intermediate horizons longer than one week, we will renounce considering such long horizons in our further investigations.

7.2.3 *Test of the time aggregation rule for heavy-tailed distribution.*

In this section, we investigate the reliability of the theoretical time aggregation rule for heavy-tailed distributions given by (6.6), which is based on the Hill estimator. We focus our study on checking this rule in the case of quantile estimates for long term horizons like one year. An obvious way to test this approach is to compare the Hill estimator (6.2) with the empirical power parameter given by (7.3), both computed on the same sample of length n ($n = 2869$ for stock indices, $n = 4173$ for foreign exchange rates), at daily frequency. Thus, for fixed time horizon ($h = 250$) and fixed probability value ($p = 0.05$) we estimate $\widehat{\alpha}^{(l)}$ and $\beta^{(l)}_{250,1}$ ($l = 1, \ldots, 500$) on 500 moving windows of length $n - 499$, i.e. we estimate $\widehat{\alpha}^{(l)}$ and $\beta^{(l)}_{250,1}$ based on the overlapping yearly returns $r^{250}_{l+249}, \ldots, r^{250}_{n+l-500}$, where $r^{250}_l = r_l + \cdots + r_{l-249}$. We plot on one graph (see Figure 10 for the DAX and Figure 11 for the USD/DEM exchange rate) the estimates

$$\left\{\left(l, 1/\widehat{\alpha}^{(l)}\right), \; l = 1, \ldots, 500\right\}$$

and

$$\left\{\left(l, \beta^{(l)}_{250,1}\right), \; l = 1, \ldots, 500\right\},$$

as well as the square-root-of-time rule (horizontal line at 0.5).

From Figure 10 one can see that for the DAX the time aggregation rule given by (6.6) is close to the empirical one. The square-root-of-time rule would lead to overestimation of the risk. In Figure 11 one can see that for foreign exchange data, using the square root of time rule leads to underestimation of the risk measure. Applying the tail rule (6.6) to heavy-tailed distributions seems to be even worse than using the square-root-of-time rule.

However, when trying to draw conclusions out of these two plots, we should keep in mind that we only had a look at the values for two data sets. Moreover, the estimates $(\widehat{\alpha}^{(i)})$ (or $(\beta^{(i)}_{250,1})$, respectively) are strongly dependent, since the $n - 998$ values of r in the center part of the time window of length n influence all estimators $(\widehat{\alpha}^{(i)})$ and $(\beta^{(i)}_{250,1})$. Hence we do not have to be very pessimistic after having seen these two graphs. For heavy-tailed distribution, we will stick to applying the scaling rule (6.6) in Section 8, where we will compare the different models.

8 Model comparison

In order to compare the models presented in Sections 3–6, it is necessary to carry out exhaustive and detailed investigations. As the performance of each model is already well known for short time horizons. For a long term issue, we first fix a horizon $h < 1$ year, for which we can use the various models. For the gap between h days and one year (kh days), we use a scaling rule. This two-step procedure, which is the key idea behind our work, is shown graphically in Figure 2.

Modelling h-day log-returns causes a first uncertainty. Scaling h-day log-returns to 1-year log-returns produces a second uncertainty. The optimal horizon h for a chosen model is the one leading to the minimal total uncertainty, in our case measured by the quality of the prediction for expected shortfall. We describe the measure we use for backtesting expected shortfall estimates in Section 8.1.2. In order to find eventually the best model among the ones we investigate, we compare the backtesting measure for all our models for several intermediate horizons h.

8.1 Backtesting.

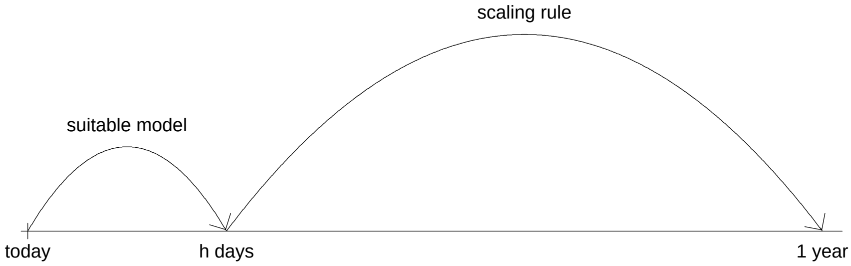

Figure 2 The two steps: first use a suitable model for a time horizon h, then scale from h days to one year.

8.1.1 *Description of the backtesting measures.*

8.1.2 *Description of the backtesting measures.*

For backtesting the forecasted expected shortfall $\widehat{\text{ES}}_t^p$, we introduce different measures. The first measure V_1^{ES} evaluates excesses below the negative of the estimated value-at-risk $\widehat{\text{VaR}}_t^p$. This is a standard method for backtesting expected shortfall estimates. In detail we proceed as follows: every model that we chose provides for each time t an estimation $\widehat{\text{ES}}_t^p$ for the one-year ahead expected shortfall ES_t^p. We build the difference between the observed one-year (k-period) return R_t^k and the negative of the estimation $\widehat{\text{ES}}_t^p$. We calculate the conditional average of these differences, where we condition on $\{R_t^k < -\widehat{\text{VaR}}_t^p\}$,

$$V_1^{\text{ES}} = \frac{\sum_{t=t_0}^{t_1} \left(R_t^k - (-\widehat{\text{ES}}_t^p) \right) \mathbf{1}_{\{R_t^k < -\widehat{\text{VaR}}_t^p\}}}{\sum_{t=t_0}^{t_1} \mathbf{1}_{\{R_t^k < -\widehat{\text{VaR}}_t^p\}}}.$$

A good estimation for expected shortfall will lead to a low absolute value of V_1^{ES}.

This first measure sticks very closely to the theoretical definition of expected shortfall. It is also dependent on the VaR estimates as we consider only values which fall below this threshold. Therefore, when analysing the outcomes given by this approach, we should combine the results with the ones given by the frequency of exceedance which we describe below.

In practice, one might primarily be interested in the loss incurred in a one in $1/p$-event, instead of getting information about the behaviour below a certain estimated value. Therefore we introduce a second measure V_2^{ES}, which evaluates values below the "1 in $1/p$-event":

$$V_2^{\text{ES}} = \frac{\sum_{t=t_0}^{t_1} D_t \, \mathbf{1}_{\{D_t < D^p\}}}{\sum_{t=t_0}^{t_1} \mathbf{1}_{\{D_t < D^p\}}},$$

where $D_t := R_t^k - (-\widehat{\text{ES}}_t^p)$ and D^p denotes the empirical p-quantile of $\{D_t\}_{t_0 \le t \le t_1}$. Note that, since $\widehat{\text{ES}}_t^p$ is an estimate on a level p, we expect D_t to be negative in less than one out of $1/p$ cases. A good estimation for expected shortfall will again lead to a low absolute value of V_2^{ES}.

The next step is to combine the two measures V_1^{ES} and V_2^{ES} as follows:

$$V^{\mathrm{ES}} = \frac{|V_1^{\mathrm{ES}}| + |V_2^{\mathrm{ES}}|}{2}.$$

This measure contains information about how well the forecasted one-year expected shortfall fits real data. It can therefore be used to backtest the quality of the models.

Finally, we introduce an additional measure that provide information about the quality of our estimators, the frequency of exceedance. It is used by the Basel Committee on Banking Supervision, which in order to encourage institutions to report their true value-at-risk numbers, devised a system in which penalties are set depending on the frequency of violations:

$$V^{\mathrm{freq}} = \frac{1}{t_1 - t_0 + 1} \sum_{t=t_0}^{t_1} \mathbf{1}_{\{R_t^k < -\widehat{\mathrm{VaR}}_t^p\}}$$

A good estimation for value-at-risk will lead to a value of V^{freq} which is close to the level p.

8.1.3 *Backtesting results.*

The methodology set up for the backtesting is described in detail for each category of data. We carry out the analysis of the results by the following two main directions. First, we extract for each model the frequency(ies) at which one should calibrate the models. Then, we compare the different methodologies considered at this previous frequency(ies). We precise that for the GARCH(1,1) model, horizons from one month upwards are missing since stationarity conditions are violated for the estimated parameters. Also, for heavy-tailed distributions we do not have enough quarterly and yearly data to estimate the tail index.

a) Foreign exchange rates, stock indices and 10-year bonds

We first use the following data sets, obtained from Datastream:

- foreign exchange rates: DEM/CHF, GBP/CHF, USD/CHF, JPY/CHF.
- stock indices: SMI, DAX, FTSE, S&P, NIKKEI.
- 10-year government bonds from CH, DE, UK, US, JP.

When trying to backtest our models, we encounter the problem that the number of yearly data is rather small to estimate model parameters and proceeding the backtesting in a significant way. We handle this problem by aggregating the backtesting results for several data sets. For each of the four foreign exchange rates we have 16 years of data (1985–2000, n=4173 daily log-returns). For both stock indices and bonds we have five samples each containing 11 years of data (1990–2000, n=2869 daily log-returns). We carry out the backtesting on each sample independently, then we aggregate the results within each of the three cases.

For each level p, each model, each intermediate horizon h and each sample, we proceed as follows:

1. We estimate the yearly forecasted expected shortfall $\widehat{\mathrm{ES}}_t^p$ on a window containing half of our data set. We use the $l = \lfloor n/2h \rfloor$ non-overlapping h-day log-returns for this estimation.
2. We compare the estimates with the real returns R_t^k.

3. We move the window by one, then we repeat steps 1 and 2 till the end of the whole data set.
4. The last step is done for the values calculated and collected in steps 1 and 2 of all data sets. We evaluate the risk measures discussed above for the pooled sample.

The results for the three data sets for $p = 1\%$ and $p = 5\%$ are presented in Tables 7, 8 and 9. Many time series models provide relatively good results for estimating the center part of distributions. The farther we go into the tail, the more distinct are the differences between their performances. Moreover, since risk managers are mainly interested in rare event cases, we will focus more on the outcomes for the 1% expected shortfall than for the 5% values. The later will be only used to confirm the reliability and the flexibility of each approach.

For foreign exchange rates (Table 7) all four models seem to perform rather well for appropriate choices of the intermediate time horizon h. More precisely, the best results are obtained when we calibrate the models either with monthly ($h = 22$) or quarterly ($h = 65$) data. We expected such behaviour since most scaling rules (RW, AR(p), HT) require some independency structure between returns and, in some cases, the normality assumption for the innovation. In Section 7, we have noticed that low frequency ($h \geq 22$) returns are closer to fulfil these hypothesis than daily or weekly returns. Concerning the models comparison, it turns out that the random walk and the heavy tailed distributions perform better than the other models. However, one should point out that if for the former approach the results are relatively stable for any frequencies and also at the 5% level, the heavy tailed approach is not reliable at other frequencies than the monthly one. This instability might be explained by the high variability of the tail index estimates due to the lack of low frequency data and/or the application of the scaling rule for high frequency returns. Although the random walk and the heavy tailed overestimate slightly the risk measures, we observe that the AR(p) model underestimates the risk measures, which can be explained through the lower variance estimates compared to the random walk variance.

For stock indices (Table 8) the same lines of analysis can be followed. Indeed, again the random walk model with $h \leq 65$ days outperforms all other models. Finally, for 10-year government bonds (Table 9) the random walk model based on daily observations provides the most reliable forecasts for one-year 1% expected shortfall.

In general for the data sets investigated, a random walk model with a constant trend and normal innovations seems to be a good choice. As these backtesting results do not provide direct information about the variability of the risk measures, we calculate, for the random walk models with normal innovations, the confidence intervals. This will be done in Section 8.4.

b) German stock data

Besides foreign exchange rates, stock indices and government bonds, another interesting class are single stocks. We now come back to German stock data, but this time not to the DAX itself, but to a data set of 22 single stocks which are part of the DAX.

As in a), there do not exist enough stationary returns per stock in order to perform the backtesting directly. In the case of German stock data, we aggregate the backtesting results for the 22 stocks. Each sample contains 23.5 years of data (n=6146 daily values). We carry out the backtesting on each sample independently, and then aggregate the results. Again, for each level p, for each model, each intermediate horizon h and each sample, we proceed as follows:

1. We estimate the yearly forecasted expected shortfall $\widehat{\mathrm{ES}}_t^p$ on a window containing half of our data set. We use the $l = \lfloor 3073/h \rfloor$ non-overlapping h-day log-returns for this estimation.
2. We compare the estimates with the real returns R_t^k.
3. We move the window by one, then we repeat steps 1 and 2 till the end of the whole data set.
4. The last step is done for the values calculated and collected in steps 1 and 2 of all data sets. We evaluate the risk measures discussed above for the pooled sample.

The results for $p = 1\%$ and $p = 5\%$ are presented in Table 10. The random walk model achieves its best performance when working directly with yearly data. For a GARCH(1,1) model a one week horizon seems to be a good choice. The results for AR(p) models do not seem to vary significantly with changing time horizon. For the heavy tailed distributions the optimal time horizon for $p = 1\%$ is one month. However, this model gives very unsatisfactory results for $p = 5\%$.

When comparing the four models with each other, we observe that heavy tailed distributions for monthly data and GARCH(1,1) models applied to weekly data provide the best forecasts, measured by their suitability to predict one-year 1% expected shortfall (backtesting measures V_1^{ES} and V_2^{ES}). As mentioned before, heavy tailed distributions for a 1-month horizon have for these 22 stocks the deficiency that they are not able to perform well on the 5% level. Moreover, as we will see in Section 8.2, for good models the backtesting results on a 5% level should indicate a slight overestimation of the risk. Our results in Table 10 show the opposite behaviour for heavy tailed distributions applied to monthly data. Because of these shortcomings of heavy tailed distributions, the GARCH(1,1) model applied to weekly data should be preferred to the heavy tailed distributions when predicting future values for stocks of the DAX.

8.2 Investigations for a simulated random walk.

In order to get a better understanding of the backtesting results presented in Section 8.1.3, we repeat the same analysis for a time series where the distribution of the log-returns is known. We saw before that random walk models with a constant trend and normal innovations are a good choice for estimating 1% expected shortfall for a one-year period. Hence we concentrate now on random walk models. After having tested the random walk models fitted to real data, we will now analyze the properties of a simulated random walk.

We proceed as follows. We simulate 1-period log-returns

$$\widetilde{r}_t^{(j)} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\widetilde{\mu}^{(j)}, \widetilde{\sigma}^{(j)}), \qquad j = 1, \ldots, 5, \quad t = h, 2h, \ldots, nh,$$

where $\widetilde{\mu}^{(j)}$ and $\widetilde{\sigma}^{(j)}$ are mean and standard deviation estimated for the 1-period (h-day) log-returns $r_t^{(j)}$ of the five stock indices SMI (j=1), DAX (j=2), FTSE (j=3), S&P (j=4) and NIKKEI (j=5) via the usual estimators

$$\widetilde{\mu}^{(j)} = \frac{1}{n} \sum_{i=1}^{n} r_{ih}^{(j)}$$

and

$$\widetilde{\sigma}^{(j)} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(r_{ih}^{(j)} - \widetilde{\mu}^{(j)})^2}.$$

We apply the method described in Section 3 to the simulated 1-period log-returns $\widetilde{r}_h^{(j)}, \widetilde{r}_{2h}^{(j)}, \dots, \widetilde{r}_{nh}^{(j)}$ to get estimates for the risk measures. For each chosen intermediate horizon h and each data set i we estimate $\widetilde{\mu}$ and $\widetilde{\sigma}$, and then calculate the corresponding one-year (k-period) value-at-risk $\widetilde{\mathrm{VaR}}^p$ and expected shortfall $\widetilde{\mathrm{ES}}^p$. Based on these estimates we can process the backtesting as described in Section 8.1. Since these results vary from simulation to simulation, we repeat the whole procedure six times. The results are displayed in Table 11, together with a recapitulation of the corresponding results for stock index data (cf. Table 8).

We observe that expected shortfall and value-at-risk tend to be overestimated. In particular, this clearly holds for a level of 5%. This is caused by the behaviour of our backtesting measures. We first investigate the frequency of exceedances V^{freq}. When evaluating this measure we calculate the portion of returns below $-\widehat{\mathrm{VaR}}^p$. This estimation for value-at-risk depends strongly on the estimated trend, i.e. on the mean of the log returns. Let Q be an estimator of a quantile of log-returns. Since the cumulative distribution function $\Phi(q)$ of a normal random variable Q is convex for $q < 0$, unbiasedness of Q (i.e. $\mathbb{E}[Q] = q$) leads to $\mathbb{E}[\Phi(Q)] \geq \Phi(q)$. In other words, the probability of a log-return being below $-\widehat{\mathrm{VaR}}^p$ is overestimated.

In absolute values, this effect is larger for a 5% VaR than for a 1% VaR. As an example, we consider the situation where the mean of a standard normal distribution is overestimated by 0.5 in the first case and underestimated by 0.5 in the second case. On average these two scenarios lead to an overestimation of the number of values which lie below the real value-at-risk of $\frac{1}{2}\Big(\Phi\big(\Phi^{-1}(5\%) + 0.5\big) + \Phi\big(\Phi^{-1}(5\%) - 0.5\big)\Big) - 5\% \approx 2.11\%$ for the 5% VaR, and $\frac{1}{2}\Big(\Phi\big(\Phi^{-1}(1\%) + 0.5\big) + \Phi\big(\Phi^{-1}(1\%) - 0.5\big)\Big) - 1\% \approx 0.81\%$ for the 1% VaR. The uncertainty in the estimated standard deviation has a similar effect. This even enlarges these differences. Since $\widehat{\mathrm{ES}}^p$ can be written as $\frac{1}{p}\int_0^p \widehat{\mathrm{VaR}}^q\, dq$, the above observation also holds for expected shortfall.

Our investigations have shown that random walks with normal innovations and a constant drift give quite good results for our data sets. To be critical, we can ask ourselves whether replacing normal innovations by heavier tailed ones, e.g. Student-t with $\nu = 4$, would lead to even better results. Having a look at our backtesting results for foreign exchange rates, stock indices and government bonds, we have to give the answer "presumably not". For most of the 15 cases (three data sets, $h = 1, 5, 22, 65, 261$), $\mathrm{VaR}^{5\%}$ is underestimated ($V^{\mathrm{freq},5\%} > 5\%$) and $\mathrm{VaR}^{1\%}$ is overestimated ($V^{\mathrm{freq},1\%} < 1\%$), see Tables 7, 8 and 9. This over- and underestimation is approximately of the same order as for the simulated time series in Table 11. Replacing normal log-returns by heavier tailed ones would even worsen the slight overestimation for the 1% level. This is due to the fact that replacing a normal distribution by a Student-t distribution with the same mean and the same 5% quantile increases the size of the estimated 1% quantile. This suggests that for the data sets used, working with normal innovations should be preferred to assuming Student-t distributed innovations in the random walk model.

8.3 Variance analysis.

8.3.1 *Description.*

In order to learn more about the behaviour of our models, we analyse the variance of forecasted yearly log-returns. We carry out this analysis in the following steps.

1. For each data set, each model and each intermediate time horizon we estimate model parameters for h-day (1-period) log-returns.
2. Based on these estimations, the one-year (k-period) variance (and the standard deviation as its square root) can be calculated.
3. To get some information about the precision of these standard deviations, we simulate 1000 times 16 years (for foreign exchange rates) and 1000 times 11 years (for stock indices and 10-year government bonds), respectively. These simulations are based on the parameters estimated in step 1. For each of the 1000 simulations we estimate the yearly standard deviation. Finally, based on these 1000 values, we calculate 95% confidence intervals for standard deviation of the model calibrated via the estimation procedure in step 1.

In the remainder of this subsection, we explain the exact proceeding within these three steps for random walk, AR(p) and GARCH(1,1) models.

a) Random walk

1. Estimate the 1-period (h-day) parameters μ and σ^2 by using the full data sets (n 1-day returns r_t^1, and $n-(h-1)$ overlapping 1-period (h-day) returns r_t):

$$\widehat{\mu} = h\,\frac{1}{n}\sum_{t=1}^{n} r_t^1,$$

$$\widehat{\sigma}^2 = \frac{1}{n-h}\sum_{t=h}^{n}(r_t-\widehat{\mu})^2.$$

2. For random walks with normal innovations, we have

$$\widetilde{\sigma}^k = \sqrt{k}\,\widehat{\sigma}.$$

3. For $l=1,\ldots,1000$: simulate $\widetilde{r}_{ih}^{(l)} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\widehat{\mu},\widehat{\sigma}^2)$ $(i=1,\ldots,m;\ m=\lfloor\frac{n}{h}\rfloor)$. Calculate the yearly standard deviation:

$$\widetilde{\sigma}^{k,(l)} = \sqrt{\frac{k}{m-1}\sum_{i=1}^{m}(\widetilde{r}_{ih}^{(l)}-\widetilde{\mu}^{(l)})^2},$$

where

$$\widetilde{\mu}^{(l)} = \frac{1}{m}\sum_{i=1}^{m}\widetilde{r}_{ih}^{(l)}.$$

This gives the 95% confidence interval $[\frac{1}{2}(\widetilde{\sigma}^{[25]}+\widetilde{\sigma}^{[26]}), \frac{1}{2}(\widetilde{\sigma}^{[975]}+\widetilde{\sigma}^{[976]})]$, where $\widetilde{\sigma}^{[l]}$ denotes the lth smallest value of the parameters $\widetilde{\sigma}^{k,(i)}$.

b) AR(p)

1. Subtract the linear trend from log-prices s_t $(t = 0, \dots, nh)$:
$$\overline{s}_t = s_t - \widehat{\mu}t, \text{ where } \widehat{\mu} = \frac{s_{nh} - s_0}{nh}.$$
Estimate the 1-period (h-day) AR(p)-parameters $\widehat{p}, \widehat{\alpha}_1, \dots, \widehat{\alpha}_{\widehat{p}}$ and $\widehat{\sigma}_\epsilon$ by applying MLE to the drift-free log-prices $\overline{s}_0, \overline{s}_h, \overline{s}_{2h}, \dots$
2. For AR($\widehat{p}$)-processes, the unconditional standard deviation of k-period (yearly) log-returns can be calculated as
$$\widetilde{\sigma}^k = \widehat{\sigma}_\epsilon \sqrt{\sum_{j=0}^{k-1} \delta_j^2},$$
where
$$\delta_j = \sum_{i=1}^{j} \widehat{\alpha}_i \delta_{j-i}, \quad j = 1, \dots, k-1,$$
$$\widehat{\alpha}_i = 0 \quad \forall i > \widehat{p},$$
$$\delta_0 = 1.$$
3. For $l = 1, \dots, 1000$: simulate AR($\widehat{p}$)-process $\widetilde{s}_{ih}^{(l)}$ $(i = 0, \dots, m;\ m = \lfloor \frac{n}{h} \rfloor)$ based on the estimates of step 1. Estimate the AR(p)-parameters $\widetilde{p}^{(l)}$, $\widetilde{\alpha}_1^{(l)}, \dots, \widetilde{\alpha}_{\widetilde{p}^{(l)}}^{(l)}$ and $\widetilde{\sigma}_\epsilon^{(l)}$ via MLE. Proceed similar to step 2 to calculate $\widetilde{\sigma}^{k,(l)}$ out of these parameters.

   This gives the 95% confidence interval $[\frac{1}{2}(\widetilde{\sigma}^{[25]} + \widetilde{\sigma}^{[26]}), \frac{1}{2}(\widetilde{\sigma}^{[975]} + \widetilde{\sigma}^{[976]})]$, where $\widetilde{\sigma}^{[l]}$ denotes the lth smallest value of the parameters $\widetilde{\sigma}^{k,(i)}$.

c) GARCH(1,1)

1. Obtain estimates $\widehat{\alpha}_0$, $\widehat{\alpha}_1$ and $\widehat{\beta}_1$ by applying QMLE to non-overlapping 1-period (h-day) log-returns of the full data sets.
2. Apply the Drost-Nijman scaling rule to get k-period (yearly) parameters $\widehat{\alpha}_{k,0}$, $\widehat{\alpha}_{k,1}$ and $\widehat{\beta}_{k,1}$. For GARCH(1,1) models the unconditional yearly standard deviation can the be estimated as
$$\widetilde{\sigma}^k = \sqrt{\frac{\widehat{\alpha}_{k,0}}{1 - (\widehat{\alpha}_{k,1} + \widehat{\beta}_{k,1})}}.$$
3. For $l = 1, \dots, 1000$: simulate GARCH(1,1)-process $\widetilde{r}_{ih}^{(l)}$ $(i = 1, \dots, m;\ m = \lfloor \frac{n}{h} \rfloor)$ based on the estimates of step 1. Use again QMLE to get estimates $\widetilde{\alpha}_0^{(l)}$, $\widetilde{\alpha}_1^{(l)}$ and $\widetilde{\beta}_1^{(l)}$. Applying Drost-Nijman gives $\widetilde{\alpha}_{k,0}^{(l)}$, $\widetilde{\alpha}_{k,1}^{(l)}$ and $\widetilde{\beta}_{k,1}^{(l)}$. For simulation number l we get the yearly standard deviation
$$\widetilde{\sigma}^{k,(l)} = \sqrt{\frac{\widetilde{\alpha}_{k,0}^{(l)}}{1 - (\widetilde{\alpha}_{k,1}^{(l)} + \widetilde{\beta}_{k,1}^{(l)})}}.$$
This gives the 95% confidence interval $[\frac{1}{2}(\widetilde{\sigma}^{[25]} + \widetilde{\sigma}^{[26]}), \frac{1}{2}(\widetilde{\sigma}^{[975]} + \widetilde{\sigma}^{[976]})]$, where $\widetilde{\sigma}^{[l]}$ denotes the lth smallest value of the parameters $\widetilde{\sigma}^{k,(i)}$.

8.3.2 *Results.*

Using the formulas given above, we get estimates for yearly standard deviations $\widetilde{\sigma}^k$, as well as simulated 95% confidence intervals. We evaluate these estimates for random walk models, AR(p) models and GARCH(1,1) models, which we apply to foreign exchange rates, stock indices and 10-year government bonds. In Figures 12–14 we plot these yearly values for intermediate time horizons varying from one day to one year.

For GARCH(1,1) models we only work with horizons up to one week, since for longer horizons stationarity conditions are violated for some of the estimated parameters. Also for the intermediate horizons of one day and one week, estimated variances for GARCH(1,1) models vary significantly. Hence some of the 95% confidence intervals are huge.

For AR(p) models, variances tend to be underestimated. Therefore also the confidence intervals are very asymmetric around the point estimate we based the simulation on. For several data sets the confidence interval for standard deviation does even not contain the value $\widetilde{\sigma}^k$, which is the true value of the standard deviation belonging to the parameters underlying the simulations.

Only for random walk models we observe a satisfying behaviour. For most data sets these variance estimates are fairly stable. As one would expect, the confidence intervals are increasing with increasing intermediate time horizon (i.e. using fewer and fewer data points).

This variance investigation confirms the conclusion of Section 8.1.3, where a random walk model with a rather small intermediate time horizon was proposed. Considering the stylised fact that daily (log) returns in a portfolio show dependencies which are not present in longer horizon returns, using an intermediate horizon between one week and one month seems to be most appropriate for modelling one-year returns.

8.4 Further investigations for random walks: confidence intervals.

In random walk models with normal innovations, we can calculate confidence intervals for one-year expected shortfall based on h-day observations. The mathematical details are given in Appendix A.2.

8.4.1 *Confidence intervals for the Swiss Market Index.*

Based on the formulas in Appendix A.2, assuming a random walk model, we calculate 95% confidence intervals for $\mathrm{VaR}^{1\%}$ and $\mathrm{ES}^{1\%}$ of the swiss market index SMI. The estimations are performed using 11 years of data (January 1, 1990 to December 31, 2000). Proceeding like this, we get one-year forecasts for December 31, 2001. The percentage loss for our model, using daily, weekly, monthly, quarterly and yearly data, is shown graphically in Figure 15. The full and the dotted lines represent 95% confidence intervals for 1% expected shortfall and 1% value-at-risk respectively. The dots visualize the corresponding point estimates. To elucidate what this percentage loss means for the absolute value of SMI, we transform these one-year point estimates and confidence intervals into predictions for December 31, 2001, see Figure 16 (observed value at the end of 2000: 8135.2; at the end of 2001: 6417.8).

For example in Figure 15, for $h = 22$ days (1 month), the point estimate for the 1% value-at-risk gives a value of 0.288. This means that with a probability of 1%, the SMI will lose 28.8% of its value or more during the year 2001. The 95% confidence interval around it goes from 0.191 to 0.388, which means that with a probability of 95% the 1% value-at-risk lies within these bounds. When we transform these percentages to values at the end of the year 2001, we get a point estimate of 5795, i.e. with a probability of 1%, the SMI will be below 5795 at the end of this one-year period. From the 95% confidence interval we can read off that with a probability of 95%, the true 99% quantile lies between 5017 and 6584.

The most evident observation in Figures 15 and 16 is the fact that the confidence intervals get larger as the intermediate time horizon h increases (i.e. as the number of values used to perform the predictions decreases). Furthermore, values for $h = 1$ year differ significantly from the ones for shorter intermediate horizons. This suggests that shorter horizons should be preferred. To figure out whether this is just a coincidence, or whether there is a clear trend as h increases, we evaluate the point estimates for all intermediate horizons h between one day and one year, and plot them in Figure 17. This plot does not confirm such a trend. But we can observe that the variation of the point estimates increases as the intermediate horizon h is increased. This suggests that a relatively small h should be chosen for estimating one-year asset risks.

8.4.2 *Confidence intervals for a simulated random walk.*

Like for the SMI data (see Section 8.4.1), we can also calculate 95% confidence intervals of $\mathrm{VaR}^{1\%}$ and $\mathrm{ES}^{1\%}$ for a simulated random walk. We estimate mean and variance for the normally distributed log-returns from SMI data. The estimations are performed by using 11 years of simulated data. The percentage yearly loss numbers for our model using daily, weekly, monthly, quarterly and yearly data, are shown graphically in Figure 18. The full and the dotted lines represent 95% confidence intervals for 1% expected shortfall and 1% value-at-risk, respectively. The dots visualize the corresponding point estimates. The slight overestimation compared to the true values ($\mathrm{ES}^{1\%} = 27.35\%$, full line; $\mathrm{VaR}^{1\%} = 23.09\%$, dotted line) stems from the fact that the average of the simulated daily log-returns was slightly smaller than the mean estimated from SMI data. Like in Figure 15, the confidence intervals get larger as the intermediate time horizon h increases.

At first glance, the estimates for simulated data seem to vary less than for the SMI. To analyse this behaviour, we evaluate the point estimates for all intermediate horizons h between one day and one year. Figure 19 shows that for simulated random walk data these estimates do actually not behave more stable than for SMI data. (For comparison: 10% in Figure 19 corresponds to 813.52 units in Figure 17.) But again, the variation of the point estimates increases as the intermediate horizon h is increased. This confirms that for a reliable prediction of one-year risks, h should not be chosen too large, as we observed already before.

9 Conclusion

This paper is concerned with the estimation of measures of market risk over long term horizons, like one year, with emphasize on the expected shortfall. We describe and test models based mostly on variance estimates like the random walk, the GARCH(1,1) and the AR(p). We propose a static approach with the heavy tailed distribution. The main motivation for choosing these models is the existence of a time aggregation rule. Since the main difficulty on measuring risk over long term is the lack of yearly data, we fit the models to higher frequency data and scale the risk estimates at the desired time horizon. The outcome of our investigations can be summarized as follows:

- Compared to short term horizon, we observe that a good estimate of the trend of returns is critical for measuring risk over a long term horizon.
- We find that for all models the best frequencies for calibration are the intermediate ones, like one week or one month. The statistical restrictions, like the sample size for estimating the model parameters and the confidence interval of the risk estimates, play an important role in this choice. Another reason is the application of scaling rules which are based on the assumptions almost fulfilled by such data.
- The random walk model performs better on average than other models. Moreover, it is the most reliable methodology since it provides equally results for both different kind of data and different confidence levels. However, like all the other models, the risk estimates for single stocks are not very satisfactory.

Finally, for one-dimensional portfolio analysis – in the sense of aggregated risk factors – we recommend to use a random walk model with a trend fitted to an intermediate time horizon like one week or one month, and to apply the square-root-of-time rule for estimating the 1% expected shortfall over one-year horizon. At last but not least it is worthy to underline the practical advantage of this approach since it is very easy to implement!

Acknowledgement The authors take pleasure in thanking Niklaus Bühlmann, John Hancock, Fabio Trojani, Paul Embrechts and Uwe Schmock for fruitful discussions, Peter Blum for providing the data, and Filip Lindskog for careful reading.

References

[1] Acerbi C. and Tasche D. (2002). *On the coherence of Expected Shortfall*, Journal of Banking and Finance 26, no. 7, 1491–1507.

[2] Ang A. and Chen J. (2000). *Asymmetric Correlations on Equity Portfolios*, Working Paper.

[3] Artzner P., Delbaen F., Eber J.M. and Heath D. (1999). *Coherent Measures of Risk*, Mathematical Finance 9, no. 3, 203–228.

[4] Barndorff-Nielsen O.E. (1977). *Exponentially Decreasing Distributions for the Logarithm of Particle Size*, Proceedings of the Royal Society London A, 353:401–419.

[5] Barndorff-Nielsen O.E. (1998). *Processes of Normal Inverse Gaussian Type*, Finance & Stochastics, 2:41–68.

[6] Borkovec M. (1999). *Large Fluctuations in Finance Models*, PhD Thesis, Munich University of Technology.

[7] Brummelhuis R. and Kaufmann R. (2003). *Estimating 10-day value-at-risk*, Working Paper.

[8] Campbell J.Y., Lo A.W. and MacKinlay A.C. (1997). *The Econometrics of Financial Markets*, Princeton University Press.

[9] Christoffersen P.F. (1997). *Evaluating Interval Forecasts*, Manuscript, Research Department, International Monetary Fund.

[10] Christoffersen P.F., Diebold F.X. and Schuermann T. (1998). *Horizon Problems and Extreme Events in Financial Risk Management*, Federal Reserve Bank of New York, 109–118.

[11] Dacorogna M.M., Müller U.A., Pictet O.V. and de Vries C.G. (1999). *Extremal Forex Returns in Extremely Large Data Sets*, Extremes, 4, 105–127.

[12] Diebold F.X. and Inoue A. (2001). *Long Memory and Regime Switching*, Journal of Econometrics, 105, 131–159.

[13] Drost F.C. and Nijman T.E. (1993). *Temporal Aggregation of GARCH Processes*, Econometrica, 61, 909–927.

[14] Duffie D. and Pan J. (1997). *An Overview of Value at Risk*, The Journal of Derivatives, Spring 1997, 7–49.

[15] Eberlein E. (1999). Application of Generalized Hyperbolic Lévy Motions to Finance. In: *Lévy Processes: Theory and Applications* (ed. Barndorff-Nielsen O.E., Mikosch T. and Resnick S.). Birkhäuser, 319–337.

[16] Eberlein E. and Keller U. (1995). *Hyperbolic Distributions in Finance*, Bernoulli, 1:281–299.

[17] Eberlein E. and Prause K. (2002) *The generalized hyperbolic model: financial derivatives and risk measures*, Mathematical Finance – Bachelier Congress 2000, Geman H., Madan D., Pliska S. and Vorst T. (eds.), Springer, 245–267.

[18] Embrechts P. (ed.) (2000). *Extremes and Integrated Risk Management*, Risk Waters Group, London.

[19] Embrechts P., Klüppelberg C. and Mikosch T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer, New York.

[20] Fama E.T. (1965). *The Behavior of Stock Market Prices*, Journal of Business, 38, 34–105.

[21] Fama E.T. and Miller M.H. (1972). *The Theory of finance*, Dryden Press.

[22] Feller W. (1971). *An Introduction to Probability Theory and its Applications, Volume II*, John Wiley, New York, 2nd edition.

[23] Gouriéroux C. (1997). *ARCH-Models and Financial Applications*, Springer Series in Statistics, Springer, New York.

[24] Hancock J., Bühlmann N. and Schmolck B. (2000). *Results of IRIS Project: Long Time-Horizon VaR*, Internal Swiss Re Report.

[25] Hull J.C. (1999). *Options, Futures and Other Derivatives*, 4th edition, Prentice-Hall.

[26] Kim J., Malz A.M. and Mina J. (1999). *LongRun Technical Document*, RiskMetrics Group, New York.

[27] Kim J. and Mina J. (2000). *ForeSight Technical Document*, RiskMetrics Group, New York.

[28] Lopez J.A. (1998). *Regulatory Evaluation of value-at-risk Models*, Federal Reserve Bank of New York, Research Paper FRBNY – No. 9710.

[29] McNeil A.J. and Frey R. (2000). *Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach*, Journal of Empirical Finance, 7, 271–300.

[30] Mikosch T. and Stărică C. (1999). *Change of structure in financial time series, long memory and the GARCH model*. Technical report, preprint available at http://www.math.ku.dk/~mikosch/preprint.html.

[31] Morgan Guaranty Trust Company (1996). *RiskMetrics Technical Document*, 4th edition.

[32] Pflug G.C. (2000), *Some remarks on the value-at-risk and the Conditional value-at-risk*, Probabilistic Constrained Optimization: Methodology and Applications, Kluwer Academic Publishers,

Dordrecht, Netherlands, 272–281.

[33] Prause K. (1999). *The Generalized hyperbolic model: Estimation, financial derivatives, and risk measures*, Dissertation, University of Freiburg.

[34] Rockafellar R.T. and Uryasev S.P. (2000). *Optimization of Conditional value-at-risk*, The Journal of Risk, Vol. 2, No. 3.

[35] Straetmans S. (1998). *Extreme Financial Returns and their Comovements*, PhD Thesis, Erasmus University Rotterdam.

[36] de Vries C.G. (1999). *Second Order Tail Effects*, EURANDOM Report 99–054.

Appendix A Proofs

A.1 Scaling rule for the kurtosis in a GARCH(1,1) model.

We give a detailed proof for the following relations between unconditional kurtosis $\widetilde{\kappa}$ (respectively $\widetilde{\kappa}^k$ for k periods) and conditional kurtosis κ (respectively κ^k for k periods) in a GARCH(1,1) model with Student-t distributed innovations:

$$\widetilde{\kappa} = \frac{1-(\alpha_1+\beta_1)^2}{1-(\alpha_1+\beta_1)^2-{\alpha_1}^2\,(\kappa-1)}\,\kappa, \tag{A.1}$$

and

$$\kappa^k = \frac{1-(\alpha_1+\beta_1)^2+\alpha_1^2}{1-(\alpha_1+\beta_1)^2+\alpha_1^2\,\widetilde{\kappa}^k}\,\widetilde{\kappa}^k. \tag{A.2}$$

We prove the two equations (A.1) and (A.2) by examining the stationary distribution. First we rewrite the GARCH(1,1) model (5.1) in the form

$$\sigma_t^2 = \alpha_0 + \alpha_1(\sigma_{t-1}\,\epsilon_{t-1})^2 + \beta_1\sigma_{t-1}^2 = \alpha_0 + \sigma_{t-1}^2(\alpha_1\epsilon_{t-1}^2+\beta_1).$$

This yields an equation for the stationary distribution:

$$\sigma^2 \stackrel{\mathrm{d}}{=} \alpha_0 + \sigma^2(\alpha_1\epsilon^2+\beta_1).$$

Taking the expected value of this expression gives

$$\mathbb{E}[\sigma^2] = \alpha_0 + \mathbb{E}[\sigma^2](\alpha_1\mathbb{E}[\epsilon^2]+\beta_1) = \alpha_0 + \mathbb{E}[\sigma^2](\alpha_1+\beta_1),$$

which can be rewritten as

$$\mathbb{E}[\sigma^2] = \frac{\alpha_0}{1-(\alpha_1+\beta_1)}.$$

For the fourth moment we get

$$\begin{aligned}
\mathbb{E}[\sigma^4] &= \mathbb{E}[(\alpha_0+\sigma^2(\alpha_1\epsilon^2+\beta_1))^2] \\
&= \alpha_0^2 + 2\alpha_0\,\mathbb{E}[\sigma^2](\alpha_1\mathbb{E}[\epsilon^2]+\beta_1) + \mathbb{E}[\sigma^4](\alpha_1^2\mathbb{E}[\epsilon^4]+2\alpha_1\beta_1\mathbb{E}[\epsilon^2]+\beta_1^2) \\
&= \alpha_0^2 + 2\alpha_0^2\frac{\alpha_1+\beta_1}{1-(\alpha_1+\beta_1)} + \mathbb{E}[\sigma^4]\big((\alpha_1+\beta_1)^2+\alpha_1^2(\mathbb{E}[\epsilon^4]-1)\big),
\end{aligned}$$

which yields

$$\mathbb{E}[\sigma^4] = \frac{\alpha_0^2(1+\alpha_1+\beta_1)}{(1-(\alpha_1+\beta_1))(1-(\alpha_1+\beta_1)^2+\alpha_1^2(1-\mathbb{E}[\epsilon^4]))}.$$

With

$$\mathbb{E}[\overline{r}^2] = \mathbb{E}[\sigma^2]\,\mathbb{E}[\epsilon^2] = \mathbb{E}[\sigma^2] = \frac{\alpha_0}{1-(\alpha_1+\beta_1)},$$

$$\mathbb{E}[\overline{r}^4] = \mathbb{E}[\sigma^4]\,\mathbb{E}[\epsilon^4] = \frac{\alpha_0^2(1+(\alpha_1+\beta_1))}{(1-(\alpha_1+\beta_1))(1-(\alpha_1+\beta_1)^2+\alpha_1^2(1-\mathbb{E}[\epsilon^4]))}\,\mathbb{E}[\epsilon^4],$$

and

$$\kappa = \frac{\mathbb{E}[\epsilon^4]}{(\mathbb{E}[\epsilon^2])^2} = \mathbb{E}[\epsilon^4],$$

we finally get (A.1):

$$\widetilde{\kappa} = \frac{\mathbb{E}[\overline{r}^4]}{(\mathbb{E}[\overline{r}^2])^2} = \frac{1-(\alpha_1+\beta_1)^2}{1-(\alpha_1+\beta_1)^2-\alpha_1^2\,(\kappa-1)}\,\kappa.$$

From this equation one can easily calculate

$$\kappa = \frac{1-(\alpha_1+\beta_1)^2+\alpha_1^2}{1-(\alpha_1+\beta_1)^2+\alpha_1^2\,\widetilde{\kappa}}\,\widetilde{\kappa}.$$

This proves (A.2), since centered k-period log-returns are assumed to fulfil equation (5.1). □

A.2 Confidence intervals for the random walk model.

For the sake of simplicity, we drop the index denoting time when deriving the confidence intervals in Section 8.4.

For fixed 1-period (h-day) parameters μ and σ, we have the estimators

$$\widehat{\mu} \sim \mathcal{N}\big(\mu, \frac{\sigma^2}{m}\big), \quad (m = \#\ h\text{-day log-returns}) \tag{A.3}$$

i.e. for fixed $\widehat{\mu}$

$$\mu \sim \mathcal{N}\big(\widehat{\mu}, \frac{\sigma^2}{m}\big),$$

and

$$(m-1)\frac{\widehat{\sigma}^2}{\sigma^2} \sim \chi^2_{m-1}. \tag{A.4}$$

From (A.4), for fixed $\widehat{\sigma}^2$, the density of the 1-period variance σ^2 can be calculated:

$$f(\sigma^2) = \frac{(m-1)^{(m-1)/2}\,\widehat{\sigma}^{m-1}}{2^{(m-1)/2}\,\Gamma(\frac{m-1}{2})\,\sigma^{m+1}}\,e^{-(m-1)\widehat{\sigma}^2/2\sigma^2}.$$

Based on this density, and based on equation (3.3), the confidence interval for expected shortfall can be derived:

$$\begin{aligned}
&\mathbb{P}\Big[-\Big(\exp\big(\mu^k+\frac{(\sigma^k)^2}{2}\big)\,\frac{\Phi(x^p-\sigma)}{p}-1\Big)\ge c\Big]\\
&=\mathbb{P}\Big[-\Big(\exp\big(k\mu+k\frac{\sigma^2}{2}\big)\,\frac{\Phi(x^p-\sqrt{k}\,\sigma)}{p}-1\Big)\ge c\Big]\\
&=\mathbb{P}\Big[k\mu\le\ln(p(1-c))-\ln\Phi(x^p-\sqrt{k}\,\sigma)-k\frac{\sigma^2}{2}\Big]\\
&=\int_0^\infty \mathbb{P}\Big[k\mu\le\ln(p(1-c))-\ln\Phi(x^p-\sqrt{k}\,\sigma)-k\frac{\sigma^2}{2}\,\Big|\,\sigma^2=s\Big]\,f(s)ds\\
&=\int_0^\infty \Phi\Big(\sqrt{\frac{m}{s}}\Big(\frac{1}{k}\ln(p(1-c))-\frac{1}{k}\ln\Phi(x^p-\sqrt{ks})-\widehat{\mu}-\frac{s}{2}\Big)\Big)\,f(s)ds\\
&=\int_0^\infty \Phi\Big(\sqrt{\frac{m}{s}}\Big(\frac{1}{k}\ln(p(1-c))-\frac{1}{k}\ln\Phi(x^p-\sqrt{ks})-\widehat{\mu}-\frac{s}{2}\Big)\Big)\,\Big(\frac{(m-1)\widehat{\sigma}^2}{2s}\Big)^{\frac{m-1}{2}}\,\frac{e^{-(m-1)\,\widehat{\sigma}^2/2s}}{s\,\Gamma(\frac{m-1}{2})}\,ds\\
&=\int_0^\infty \Phi\Big(\sqrt{\frac{2m}{(m-1)\widehat{\sigma}^2u}}\Big(\frac{1}{k}\ln(p(1-c))-\frac{1}{k}\ln\Phi(x^p-\sqrt{\frac{k(m-1)\widehat{\sigma}^2u}{2}})-\widehat{\mu}-\frac{(m-1)\widehat{\sigma}^2u}{4}\Big)\Big)\frac{e^{-1/u}}{u^{(m+1)/2}\,\Gamma(\frac{m-1}{2})}du\\
&=: g(c).
\end{aligned}$$

The two sided $(1-\alpha)$ confidence interval for expected shortfall is $[g^{-1}(1-\frac{\alpha}{2}), g^{-1}(\frac{\alpha}{2})]$. This interval can be calculated numerically.

In the same way, based on equation (3.2), we can derive the confidence interval for value-at-risk:

$$
\begin{aligned}
&\mathbb{P}[-\Big(\exp\big(\mu^k+\sigma^k x^p\big)-1\Big)\geq c] \\
&\quad= \mathbb{P}[-\Big(\exp\big(k\mu+\sqrt{k}\sigma x^p\big)-1\Big)\geq c] \\
&\quad= \mathbb{P}[k\mu \leq \ln(1-c)-\sqrt{k}\sigma x^p] \\
&\quad= \int_0^\infty \mathbb{P}[k\mu \leq \ln(1-c)-\sqrt{k}\sigma x^p \,|\, \sigma^2=s]\, f(s)ds \\
&\quad= \int_0^\infty \Phi\bigg(\sqrt{\frac{m}{s}}\Big(\frac{1}{k}\ln(1-c)-\sqrt{\frac{s}{k}}x^p-\widehat{\mu}\Big)\bigg)\, f(s)ds \\
&\quad= \int_0^\infty \Phi\bigg(\sqrt{\frac{m}{s}}\Big(\frac{1}{k}\ln(1-c)-\sqrt{\frac{s}{k}}x^p-\widehat{\mu}\Big)\bigg)\, \Big(\frac{(m-1)\widehat{\sigma}^2}{2s}\Big)^{\frac{m-1}{2}}\, \frac{e^{-(m-1)\,\widehat{\sigma}^2/2s}}{s\,\Gamma(\frac{m-1}{2})}\, ds \\
&\quad= \int_0^\infty \Phi\bigg(\sqrt{\frac{2m}{(m-1)\widehat{\sigma}^2 u}}\Big(\frac{1}{k}\ln(1-c)-\sqrt{\frac{(m-1)\widehat{\sigma}^2 u}{2k}}x^p-\widehat{\mu}\Big)\bigg)\frac{e^{-1/u}}{u^{(m+1)/2}\,\Gamma(\frac{m-1}{2})}\, du \\
&\quad= \int_0^\infty \Phi\bigg(\sqrt{\frac{2m}{(m-1)\widehat{\sigma}^2 u}}\Big(\frac{1}{k}\ln(1-c)-\widehat{\mu}\Big)-\sqrt{\frac{m}{k}}x^p\bigg)\frac{e^{-1/u}}{u^{(m+1)/2}\,\Gamma(\frac{m-1}{2})}\, du =: h(c).
\end{aligned}
$$

The two sided $(1-\alpha)$ confidence interval for value-at-risk is $[h^{-1}(1-\frac{\alpha}{2}), h^{-1}(\frac{\alpha}{2})]$, which can also be calculated numerically.

Appendix B Tables

Data	Frequency	Length	Mean	Variance	Skewness	Kurtosis
DEM/CHF	1	4173	-0.00001	0.0001	-0.325	7.14
	5	834	-0.00007	0.0001	-0.315	4.36
	10	417	-0.00014	0.0001	-0.298	3.49
	20	208	-0.00033	0.0001	-0.524	3.90
	30	139	-0.00041	0.0002	-0.385	3.46
GBP/CHF	1	4173	-0.00005	0.0001	-0.613	11.80
	5	834	-0.00025	0.0002	-0.686	9.37
	10	417	-0.00049	0.0003	-0.862	7.66
	20	208	-0.00093	0.0006	-0.657	4.56
	30	139	-0.00147	0.0011	-0.758	5.50
USD/CHF	1	4173	-0.00011	0.0001	-0.271	5.47
	5	834	-0.00055	0.0003	-0.112	3.84
	10	417	-0.00109	0.0006	-0.042	3.37
	20	208	-0.00201	0.0013	-0.111	3.00
	30	139	-0.00328	0.0019	-0.146	3.16
JPY/CHF	1	4173	0.00008	0.0001	0.274	6.04
	5	834	0.00042	0.0002	0.586	5.77
	10	417	0.00085	0.0005	0.629	5.18
	20	208	0.00188	0.0010	0.226	3.32
	30	139	0.00255	0.0015	0.017	3.18

Table 1 Figures of foreign exchange rates: DEM/CHF, GBP/CHF, USD/CHF and JPY/CHF from 01.01.1985 to 29.12.2000 considered at different frequencies. (We recall that we work with h-day log-returns $r_t = \log(\frac{S_t}{S_{t-h}})$.)

Data	Frequency	Length	Mean	Variance	Skewness	Kurtosis
SMI	1	2779	0.0005	0.0001	-0.374	8.42
	5	555	0.0027	0.0006	-0.449	4.41
	10	277	0.0054	0.0011	-0.381	4.94
	20	138	0.0192	0.0022	-0.525	4.87
	30	92	0.0163	0.0029	-0.412	4.39
DAX	1	2807	0.0004	0.0002	-0.401	7.54
	5	561	0.0018	0.0008	-0.368	4.69
	10	280	0.0038	0.0016	-0.175	4.54
	20	140	0.0075	0.0035	-0.354	4.95
	30	93	0.0114	0.0047	-0.495	3.75
FTSE	1	2786	0.0003	0.0001	0.013	4.82
	5	557	0.0017	0.0004	-0.086	4.42
	10	278	0.0034	0.0008	0.076	3.71
	20	139	0.0068	0.0014	-0.155	2.98
	30	92	0.0105	0.0025	-0.103	3.49
S&P	1	2817	0.0005	0.0001	-0.294	7.82
	5	563	0.0023	0.0005	-0.792	7.11
	10	281	0.0046	0.0007	-0.358	4.67
	20	140	0.0097	0.0016	-0.363	4.91
	30	93	0.0144	0.0022	-0.931	9.01
NIKKEI	1	2802	-0.0002	0.0002	0.042	8.15
	5	560	-0.0012	0.0009	0.091	5.01
	10	280	-0.0024	0.0019	0.385	6.69
	20	140	-0.005	0.0036	0.111	3.44
	30	93	-0.0063	0.0049	0.339	3.13

Table 2 Figures of stock indices: SMI, DAX, FTSE, S&P and NIKKEI from 01.01.1990 to 29.12.2000 considered at different frequencies.

Data	Frequency	Length	Mean	Variance	Skewness	Kurtosis
CH-B10y	1	2779	-0.0002	0.0001	0.279	9.73
	5	555	-0.0007	0.0003	0.005	4.61
	10	277	-0.0014	0.0008	0.146	4.21
	20	138	-0.0027	0.0016	0.055	3.29
	30	92	-0.0041	0.0029	-0.126	3.37
DE-B10y	1	2807	-0.0001	0.0001	0.338	5.36
	5	561	-0.0007	0.0003	0.839	4.93
	10	280	-0.0013	0.0005	0.694	4.58
	20	140	-0.0027	0.0011	0.688	3.91
	30	93	-0.0037	0.0021	0.851	3.87
GB-B10y	1	2786	-0.0002	0.0001	0.296	6.96
	5	557	-0.0012	0.0003	0.363	5.29
	10	278	-0.0025	0.0006	0.065	3.22
	20	139	-0.0051	0.0011	0.050	3.39
	30	92	-0.0074	0.0018	0.279	3.05
US-B10y	1	2817	-0.0001	0.0001	0.232	5.39
	5	563	-0.0007	0.0004	0.269	4.26
	10	281	-0.0012	0.0007	0.381	3.41
	20	140	-0.0020	0.0014	0.423	3.03
	30	93	-0.0027	0.0022	0.443	2.91
JP-B10y	1	2869	-0.0004	0.0002	0.186	11.14
	5	573	-0.0022	0.0011	0.131	7.98
	10	286	-0.0043	0.0022	0.177	8.63
	20	143	-0.0091	0.0053	0.596	5.75
	30	95	-0.0131	0.0065	0.957	5.51

Table 3 Figures of 10-year government bonds for Switzerland, Germany, Great Britain, USA and Japan from 01.01.1990 to 29.12.2000 considered at different frequencies.

Deutscher Aktienindex (DAX)		
$h = 5$ days (1 week)	Drost-Nijman	$\sigma_t^2 = 6.586 \cdot 10^{-5} + 0.10485\, \overline{r}_{t-5}^2 + 0.7924\, \sigma_{t-5}^2$
	direct estimation	$\sigma_t^2 = 1.751 \cdot 10^{-5} + 0.09731\, \overline{r}_{t-5}^2 + 0.8710\, \sigma_{t-5}^2$
$h = 20$ days (1 month)	Drost-Nijman	$\sigma_t^2 = 9.023 \cdot 10^{-4} + 0.09640\, \overline{r}_{t-20}^2 + 0.5519\, \sigma_{t-20}^2$
	direct estimation	$\sigma_t^2 = 5.740 \cdot 10^{-5} + 0.09217\, \overline{r}_{t-20}^2 + 0.8847\, \sigma_{t-20}^2$
$h = 80$ days (4 months)	Drost-Nijman	$\sigma_t^2 = 8.449 \cdot 10^{-3} + 0.04016\, \overline{r}_{t-80}^2 + 0.1364\, \sigma_{t-80}^2$
	direct estimation	$\sigma_t^2 = 2.907 \cdot 10^{-3} + 0.11495\, \overline{r}_{t-80}^2 + 0.6095\, \sigma_{t-80}^2$
$h = 261$ days (1 year)	Drost-Nijman	$\sigma_t^2 = 3.336 \cdot 10^{-2} + 0.00665\, \overline{r}_{t-261}^2 - 0.0032\, \sigma_{t-261}^2$
	direct estimation	$\sigma_t^2 = 1.696 \cdot 10^{-2} - 0.02506\, \overline{r}_{t-261}^2 + 0.4420\, \sigma_{t-261}^2$
$h \to \infty$	Drost-Nijman	$\sigma_t^2 = 1.283 \cdot 10^{-4}\, h$

Table 4 Parameter estimates for the DAX.

USD/DEM exchange rate		
$h = 5$ days (1 week)	Drost-Nijman	$\sigma_t^2 = 1.097 \cdot 10^{-5} + 0.06977\, \overline{r}_{t-5}^2 + 0.8840\, \sigma_{t-5}^2$
	direct estimation	$\sigma_t^2 = 1.035 \cdot 10^{-5} + 0.08140\, \overline{r}_{t-5}^2 + 0.8785\, \sigma_{t-5}^2$
$h = 20$ days (1 month)	Drost-Nijman	$\sigma_t^2 = 1.637 \cdot 10^{-4} + 0.08110\, \overline{r}_{t-20}^2 + 0.7463\, \sigma_{t-20}^2$
	direct estimation	$\sigma_t^2 = 1.419 \cdot 10^{-4} + 0.16233\, \overline{r}_{t-20}^2 + 0.7152\, \sigma_{t-20}^2$
$h = 80$ days (4 months)	Drost-Nijman	$\sigma_t^2 = 2.016 \cdot 10^{-3} + 0.05766\, \overline{r}_{t-80}^2 + 0.4109\, \sigma_{t-80}^2$
	direct estimation	$\sigma_t^2 = 5.096 \cdot 10^{-3} + 0.07089\, \overline{r}_{t-80}^2 - 0.0867\, \sigma_{t-80}^2$
$h = 261$ days (1 year)	Drost-Nijman	$\sigma_t^2 = 1.133 \cdot 10^{-2} + 0.01835\, \overline{r}_{t-261}^2 + 0.0660\, \sigma_{t-261}^2$
	direct estimation	$\sigma_t^2 = 9.554 \cdot 10^{-3} - 0.15638\, \overline{r}_{t-261}^2 + 0.5892\, \sigma_{t-261}^2$
$h \to \infty$	Drost-Nijman	$\sigma_t^2 = 4.742 \cdot 10^{-5}\, h$

Table 5 Parameter estimates for USD/DEM exchange rates.

simulated GARCH(1,1) with $\alpha_0 = 2 \cdot 10^{-5}, \alpha_1 = 0.08, \beta_1 = 0.90$		
$h = 5$ days (1 week)	Drost-Nijman	$\sigma_t^2 = 4.804 \cdot 10^{-5} + 0.09191\, \overline{r}_{t-5}^2 + 0.8120\, \sigma_{t-5}^2$
	direct estimation	$\sigma_t^2 = 5.418 \cdot 10^{-5} + 0.09633\, \overline{r}_{t-5}^2 + 0.7737\, \sigma_{t-5}^2$
$h = 20$ days (1 month)	Drost-Nijman	$\sigma_t^2 = 6.648 \cdot 10^{-4} + 0.08562\, \overline{r}_{t-20}^2 + 0.5820\, \sigma_{t-20}^2$
	direct estimation	$\sigma_t^2 = 6.213 \cdot 10^{-4} + 0.12237\, \overline{r}_{t-20}^2 + 0.4765\, \sigma_{t-20}^2$
$h = 80$ days (4 months)	Drost-Nijman	$\sigma_t^2 = 6.411 \cdot 10^{-3} + 0.03696\, \overline{r}_{t-80}^2 + 0.1617\, \sigma_{t-80}^2$
	direct estimation	$\sigma_t^2 = 4.278 \cdot 10^{-4} + 0.11777\, \overline{r}_{t-80}^2 + 0.8120\, \sigma_{t-80}^2$
$h = 261$ days (1 year)	Drost-Nijman	$\sigma_t^2 = 2.597 \cdot 10^{-2} + 0.00626\, \overline{r}_{t-261}^2 - 0.0011\, \sigma_{t-261}^2$
	direct estimation	$\sigma_t^2 = -1.455 \cdot 10^{-3} - 0.11527\, \overline{r}_{t-261}^2 + 1.1950\, \sigma_{t-261}^2$
$h \to \infty$	Drost-Nijman	$\sigma_t^2 = 1.000 \cdot 10^{-4}\, h$

Table 6 Parameter estimates for a simulated GARCH(1,1) process.

Model	Freq days	$ES^{1\%}$ V^{ES}	$ES^{1\%}$ V_1^{ES}	$ES^{1\%}$ V_2^{ES}	VaR V^{freq}	$ES^{5\%}$ V^{ES}	$ES^{5\%}$ V_1^{ES}	$ES^{5\%}$ V_2^{ES}	VaR V^{freq}
Optimal		0%	0%	0%	1%	0%	0%	0%	5%
Random Walk	1	1.4%	1.1%	1.7%	0.4%	1.0%	0.8%	-1.3%	8.6%
	5	1.1%	0.8%	1.4%	0.5%	1.1%	1.0%	-1.3%	8.7%
	22	1.0%	0.7%	1.3%	0.5%	1.2%	1.2%	-1.2%	8.1%
	65	0.9%	0.5%	1.4%	0.5%	1.1%	1.1%	-1.1%	7.1%
	261	0.8%	-0.4%	-1.1%	1.6%	1.7%	0.1%	-3.3%	9.4%
GARCH(1,1)	1	N/A	N/A	3.9%	0.0%	1.9%	3.7%	0.0%	7.3%
	5	2.2%	1.9%	2.6%	0.3%	1.5%	2.0%	-1.1%	10.1%
AR(p)	1	2.3%	-0.7%	-3.9%	6.6%	3.9%	-1.1%	-6.8%	19.5%
	5	2.4%	-0.9%	-4.0%	5.9%	3.9%	-0.9%	-6.9%	19.1%
	22	2.3%	-0.8%	-3.8%	5.9%	3.8%	-0.8%	-6.8%	18.8%
	65	2.1%	-0.7%	-3.4%	5.8%	3.6%	-0.8%	-6.4%	18.3%
	261	5.2%	-1.9%	-8.6%	10.7%	6.9%	-2.3%	-11.4%	23.4%
Heavy-tailed Distribution	1	5.1%	-1.2%	-8.9%	15.2%	9.4%	-3.1%	-15.8%	41.4%
	5	3.7%	0.2%	-7.1%	9.3%	7.8%	-1.8%	-13.8	35.8%
	22	0.8%	1.6%	0.1%	1.4%	4.8%	0.1%	-9.4%	19.8%

Table 7 Backtesting results for the levels $p = 1\%$ and $p = 5\%$, using exchange rate data: DEM/CHF, GBP/CHF, USD/CHF, JPY/CHF. Measures for the value-at-risk and expected shortfall are evaluated for a one-year horizon. For the GARCH(1,1) model, horizons from one month upwards are missing since stationarity conditions are violated for the estimated parameters. For heavy-tailed distributions we do not have enough quarterly and yearly data to estimate the tail index with the Hill estimator.

Model	Freq days	$\mathrm{ES}^{1\%}$ V^{ES}	V_1^{ES}	V_2^{ES}	VaR V^{freq}	$\mathrm{ES}^{5\%}$ V^{ES}	V_1^{ES}	V_2^{ES}	VaR V^{freq}
Optimal		0%	0%	0%	1%	0%	0%	0%	5%
Random Walk	1	0.8%	0.3%	1.3%	0.8%	3.5%	0.0%	-7.0%	8.3%
	5	1.2%	0.5%	1.9%	0.7%	3.2%	0.0%	-6.4%	8.1%
	22	0.7%	0.2%	1.1%	0.8%	3.7%	-0.2%	-7.1%	8.3%
	65	1.3%	-1.2%	-1.3%	1.0%	4.7%	-0.8%	-8.5%	8.6%
	261	10.5%	-6.0%	-15.0%	2.5%	11.0%	-4.0%	-18.0%	9.2%
GARCH(1,1)	1	0.6%	0.2%	-1.1%	1.3%	5.4%	-0.3%	-10.5%	8.8%
	5	3.7%	2.6%	4.9%	0.5%	3.1%	1.6%	-4.6%	7.6%
AR(p)	1	6.3%	-3.2%	-9.3%	3.3%	8.5%	-3.5%	-13.5%	10.6%
	5	6.4%	-3.2%	-9.6%	3.4%	8.5%	-3.4%	-13.7%	11.0%
	22	7.1%	-3.8%	-10.4%	3.2%	8.8%	-3.2%	-14.4%	11.2%
	65	8.8%	-4.1%	-13.4%	3.8%	9.9%	-3.3%	-16.6%	12.3%
	261	13.7%	-6.4%	-21.0%	12.4%	14.7%	-6.8%	-22.6%	20.9%
Heavy-tailed Distribution	1	3.0%	4.1%	1.9%	2.0%	2.5%	1.5%	3.5%	3.6%
	5	2.4%	1.8%	2.9%	0.8%	4.5%	2.4%	6.7	2.3%
	22	1.7%	-0.5%	2.9%	0.5%	8.4%	3.0%	13.8%	0.9%

Table 8 Backtesting results for the levels $p = 1\%$ and $p = 5\%$, using stock index data: SMI, DAX, FTSE, S&P, NIKKEI. Measures for the value-at-risk and expected shortfall are evaluated for a one-year horizon.

Model	Freq days	$\text{ES}^{1\%}$ V^{ES}	V_1^{ES}	V_2^{ES}	VaR V^{freq}	$\text{ES}^{5\%}$ V^{ES}	V_1^{ES}	V_2^{ES}	VaR V^{freq}
Optimal		0%	0%	0%	1%	0%	0%	0%	5%
Random Walk	1	1.0%	0.3%	1.8%	0.6%	1.9%	0.5%	-3.3%	7.3%
	5	1.8%	0.5%	3.2%	0.4%	1.6%	0.8%	-2.4%	7.1%
	22	2.4%	-0.4%	4.4%	0.2%	1.2%	1.3%	-1.1%	6.7%
	65	3.6%	1.1%	6.1%	0.1%	1.0%	1.6%	0.5%	5.6%
	261	4.1%	-4.7%	-3.4%	0.9%	5.2%	-2.8%	-7.7%	6.0%
GARCH(1,1)	1	6.1%	-1.8%	10.4%	0.0%	4.8%	4.4%	5.2%	4.4%
	5	10.7%	11.8%	9.5%	0.1%	2.3%	2.6%	2.0%	5.1%
AR(p)	1	5.9%	-2.4%	-9.3%	3.4%	7.8%	-2.1%	-13.5%	12.4%
	5	5.8%	-2.6%	-9.0%	3.0%	7.6%	-1.9%	-13.2%	11.9%
	22	5.7%	-2.8%	-8.5%	2.8%	7.4%	-1.9%	-12.8%	11.9%
	65	6.9%	-3.6%	-10.2%	3.4%	8.2%	-2.3%	-14.2%	13.6%
	261	13.1%	-4.9%	-21.2%	13.5%	15.0%	-5.9%	-24.1%	23.8%
Heavy-tailed Distribution	1	11.4%	-2.2%	-20.5%	35.0%	9.4%	-3.1%	-15.8%	41.4%
	5	8.4%	-1.2%	-15.6%	25.2%	7.8%	-1.8%	-13.8	35.8%
	22	7.6%	-1.4%	-13.7%	11.7%	4.8%	0.1%	-9.4%	19.8%

Table 9 Backtesting results for the levels $p = 1\%$ and $p = 5\%$, using 10-year government bond data: CH, DE, UK, US, JP. Measures for the value-at-risk and expected shortfall are evaluated for a one-year horizon.

Model	Freq days	$\mathrm{ES}^{1\%}$ V^{ES}	V_1^{ES}	V_2^{ES}	VaR V^{freq}	$\mathrm{ES}^{5\%}$ V^{ES}	V_1^{ES}	V_2^{ES}	VaR V^{freq}
Optimal		0%	0%	0%	1%	0%	0%	0%	5%
Random Walk	1	10.3%	-5.2%	-15.4%	3.3%	6.8%	-4.2%	-9.4%	8.3%
	5	9.6%	-4.7%	-14.6%	3.3%	6.5%	-4.0%	-9.1%	8.3%
	22	8.3%	-4.1%	-12.5%	2.9%	5.8%	-3.7%	-7.9%	7.7%
	65	9.0%	-4.5%	-13.6%	3.0%	6.1%	-4.0%	-8.2%	7.5%
	261	4.8%	-2.6%	-7.0%	2.1%	3.3%	-2.5%	-4.1%	6.0%
GARCH(1,1)	1	6.1%	-2.6%	-9.5%	2.3%	4.2%	-2.2%	-6.1%	7.5%
	5	2.6%	-1.4%	-3.8%	1.4%	1.3%	-0.2%	-2.4%	6.5%
AR(p)	1	8.8%	-3.8%	-13.9%	4.4%	6.8%	-3.3%	-10.2%	11.2%
	5	8.6%	-3.8%	-13.3%	4.0%	6.3%	-3.0%	-9.6%	10.8%
	22	7.7%	-3.6%	-11.8%	3.4%	5.3%	-2.5%	-8.2%	9.9%
	65	7.3%	-3.3%	-11.2%	3.5%	5.2%	-2.4%	-8.1%	10.0%
	261	9.8%	-4.3%	-15.3%	3.2%	5.8%	-2.8%	-8.9%	8.9%
Heavy-tailed Distribution	1	6.1%	1.3%	-10.8%	5.2%	2.8%	2.0%	-3.5%	10.1%
	5	6.2%	-0.3%	-12.1%	4.7%	2.6%	0.5%	-4.7%	8.9%
	22	1.7%	2.6%	0.8%	1.1%	7.2%	5.7%	8.6%	3.0%

Table 10 Backtesting results for the levels $p = 1\%$ and $p = 5\%$ for the DAX. Measures for the value-at-risk and expected shortfall are evaluated for a one-year horizon. For the GARCH(1,1) model, horizons from one month upwards are missing since stationarity conditions are violated for the estimated parameters. For heavy-tailed distributions we do not have enough quarterly and yearly data to estimate the tail index with the Hill estimator.

Data	Freq days	ES$^{1\%}$ V^{ES}	ES$^{1\%}$ V_1^{ES}	ES$^{1\%}$ V_2^{ES}	VaR V^{freq}	ES$^{5\%}$ V^{ES}	ES$^{5\%}$ V_1^{ES}	ES$^{5\%}$ V_2^{ES}	VaR V^{freq}
Optimal		0%	0%	0%	1%	0%	0%	0%	5%
Simulation 1	1	0.9%	1.2%	0.6%	1.2%	4.4%	0.4%	-8.5%	9.4%
	5	1.8%	1.9%	1.7%	1.1%	4.1%	0.6%	-7.6%	9.1%
	22	0.4%	0.7%	0.1%	1.2%	4.6%	0.1%	-9.0%	9.2%
	65	2.5%	-1.1%	-3.8%	1.9%	6.6%	-0.9%	-12.3%	11.1%
	261	12.0%	-5.2%	-18.8%	7.1%	14.9%	-5.3%	-24.6%	16.5%
Simulation 2	1	1.3%	0.9%	-1.7%	2.9%	5.6%	-1.1%	-10.0%	12.7%
	5	1.4%	0.4%	-2.4%	3.0%	6.0%	-1.3%	-10.7%	13.1%
	22	2.1%	-0.3%	-3.9%	3.2%	6.9%	-1.8%	-11.9%	14.0%
	65	2.4%	-0.2%	-4.5%	3.9%	7.1%	-2.0%	-12.3%	15.0%
	261	9.6%	-4.4%	-14.8%	6.3%	12.7%	-3.8%	-21.6%	18.2%
Simulation 3	1	4.6%	3.3%	6.0%	0.2%	2.0%	2.1%	-1.9%	7.6%
	5	4.4%	3.1%	5.7%	0.3%	2.0%	2.1%	-2.0%	7.6%
	22	2.8%	1.4%	4.3%	0.4%	2.3%	1.5%	-3.1%	7.2%
	65	3.0%	2.0%	4.0%	0.5%	2.6%	1.8%	-3.4%	8.1%
	261	10.1%	-4.3%	-15.8%	3.9%	11.5%	-2.0%	-20.9%	15.6%
Simulation 4	1	2.5%	1.7%	3.4%	0.6%	3.3%	1.5%	-5.2%	11.0%
	5	3.1%	1.9%	4.3%	0.4%	2.9%	1.6%	-4.1%	10.9%
	22	0.5%	0.1%	0.8%	0.9%	4.1%	0.9%	-7.2%	12.5%
	65	1.1%	0.8%	-1.4%	1.0%	4.9%	0.4%	-9.3%	14.6%
	261	6.6%	-2.7%	-10.6%	5.6%	9.2%	-3.3%	-15.1%	12.1%
Simulation 5	1	1.9%	0.0%	-3.7%	4.5%	7.8%	-3.8%	-11.8%	12.4%
	5	2.6%	-0.5%	-4.8%	4.5%	8.4%	-4.1%	-12.6%	12.3%
	22	3.8%	-1.4%	-6.1%	4.3%	9.0%	-4.0%	-13.9%	12.6%
	65	5.7%	-2.4%	-9.0%	4.5%	10.4%	-4.3%	-16.4%	12.9%
	261	13.8%	-7.5%	-20.1%	6.7%	16.5%	-7.2%	-25.8%	16.2%
Simulation 6	1	0.7%	1.2%	0.2%	1.5%	3.6%	-0.1%	-7.0%	10.5%
	5	0.8%	0.8%	-0.8%	1.8%	4.1%	-0.4%	-7.8%	10.8%
	22	0.6%	1.0%	0.3%	1.3%	3.5%	0.2%	-6.9%	10.8%
	65	2.7%	2.3%	3.0%	0.7%	2.5%	0.6%	-4.4%	10.2%
	261	7.1%	-4.0%	-10.2%	4.3%	8.4%	-2.7%	-14.0%	13.8%
Stock Indices	1	0.8%	0.3%	1.3%	0.8%	3.5%	0.0%	-7.0%	8.3%
	5	1.2%	0.5%	1.9%	0.7%	3.2%	0.0%	-6.4%	8.1%
	22	0.7%	0.2%	1.1%	0.8%	3.7%	-0.2%	-7.1%	8.3%
	65	1.3%	-1.2%	-1.3%	1.0%	4.7%	-0.8%	-8.5%	8.6%
	261	10.5%	-6.0%	-15.0%	2.5%	11.0%	-4.0%	-18.0%	9.2%

Table 11 Backtesting results for a random walk model for the levels $p = 1\%$ and $p = 5\%$, using simulated random walk data and real stock indices, respectively. Measures for the value-at-risk and expected shortfall are evaluated for a one-year horizon.

Appendix C Graphs

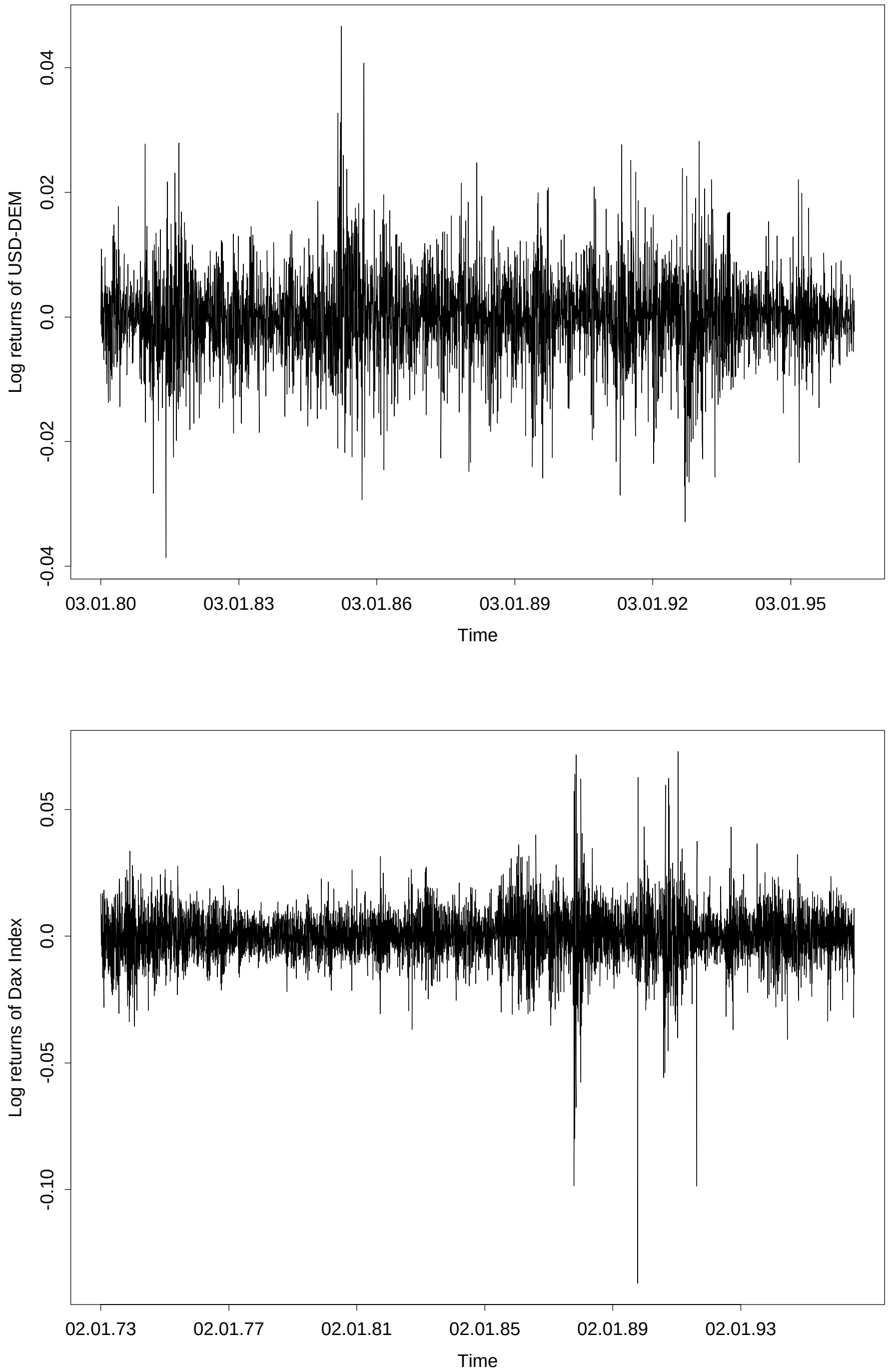

Figure 3 Daily log-returns of USD/DEM exchange rates (top) and the DAX (bottom).

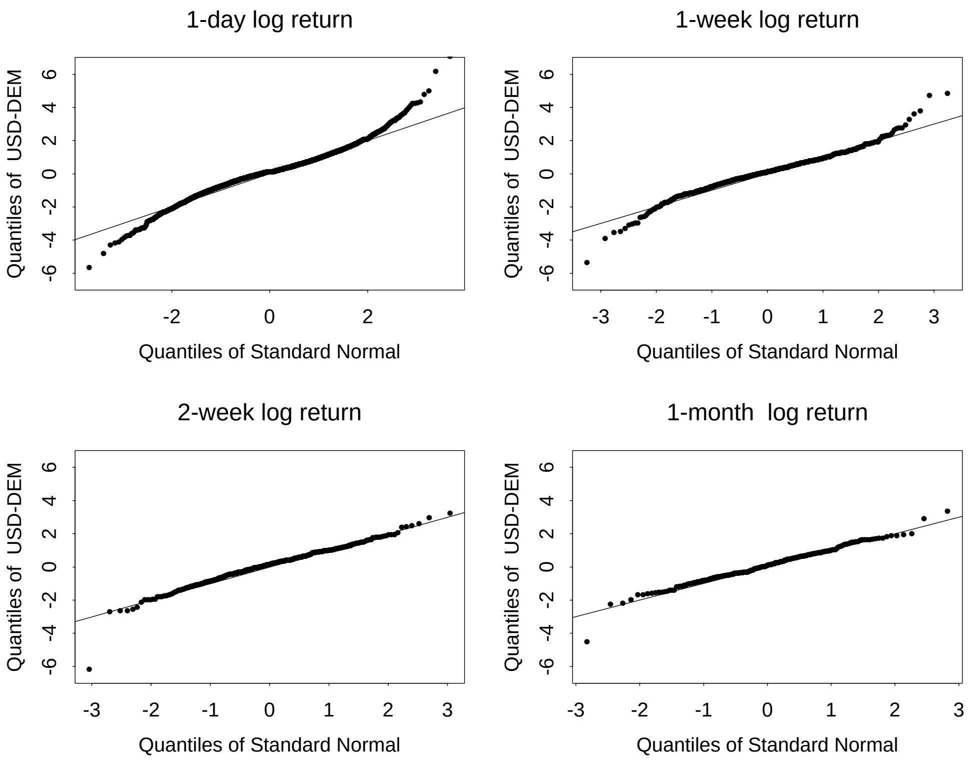

Figure 4 QQ-plots of USD/DEM exchange rates for different frequencies.

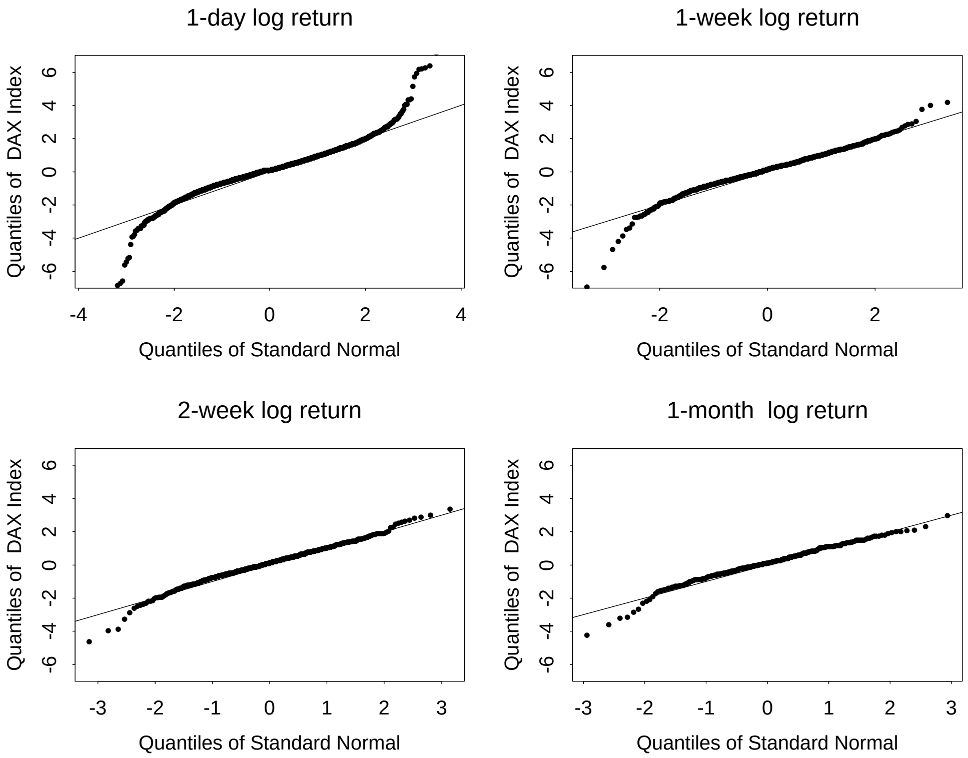

Figure 5 QQ-plots of the DAX for different frequencies.

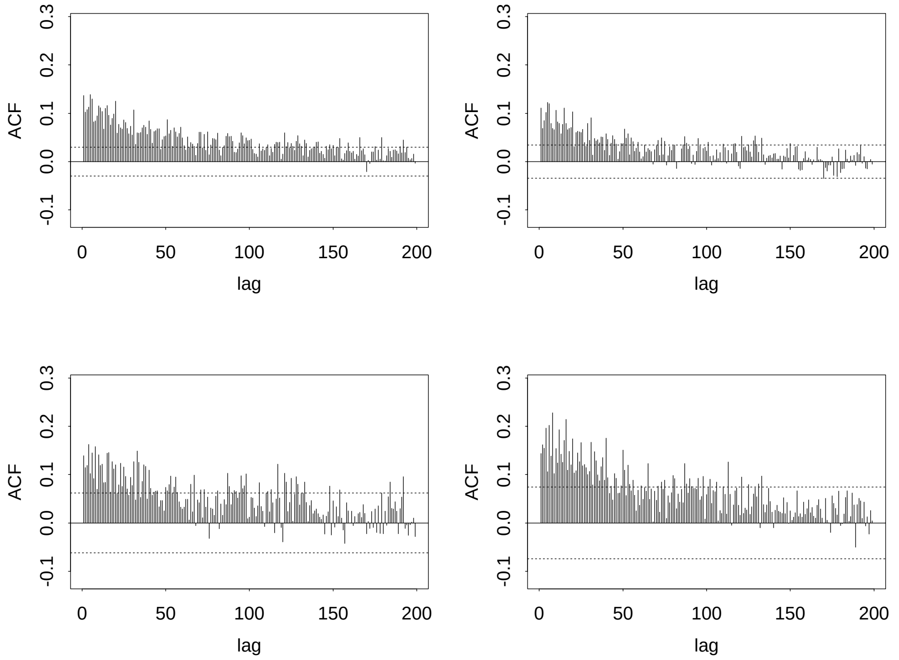

Figure 6 Sample autocorrelation functions of daily, absolute returns of USD/DEM exchange rates for different periods. Upper left: whole sample, upper right: 02.01.1982 - 23.10.1987, lower left: 29.12.1982 - 22.01.1990, lower right: 29.12.1983 - 24.03.1992.

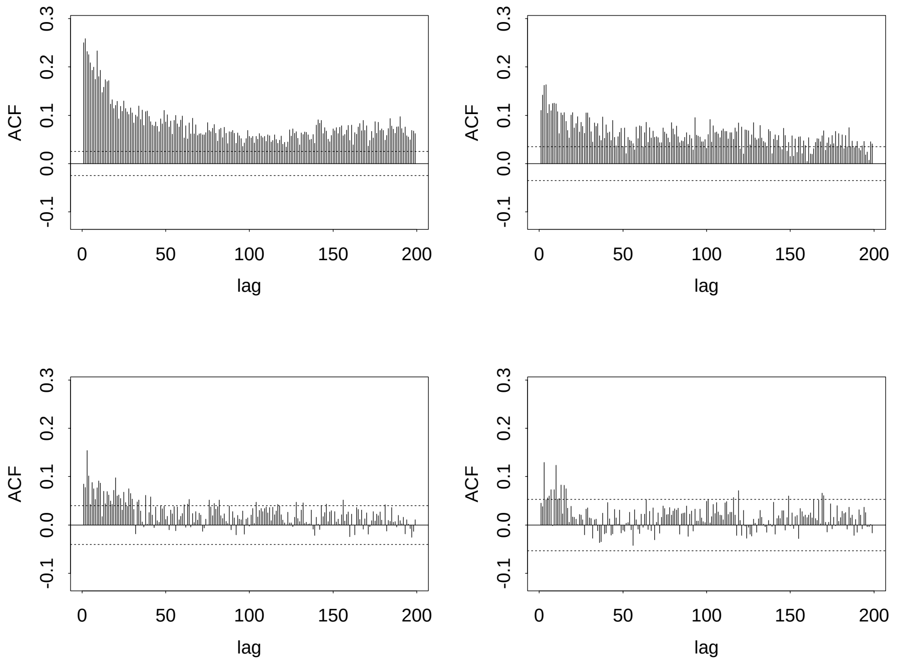

Figure 7 Sample autocorrelation functions of daily, absolute returns of the DAX for different periods. Upper left: whole sample, upper right: 02.01.1973 - 23.10.1990, lower left: 29.12.1975 - 22.01.1985, lower right: 29.12.1975 - 24.03.1981.

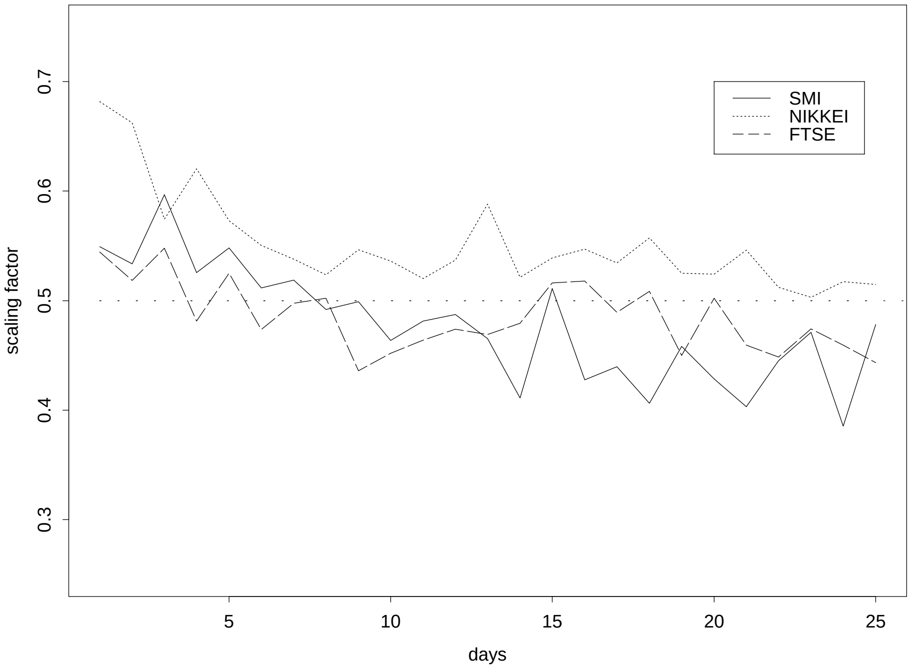

Figure 8 Empirical time aggregation rule (non-overlapping case) – power parameter plot with respect to quantile frequency for some stock indices with $p = 0.01$. The dashed line at 0.5 corresponds to the power parameter when applying the square-root-of-time rule.

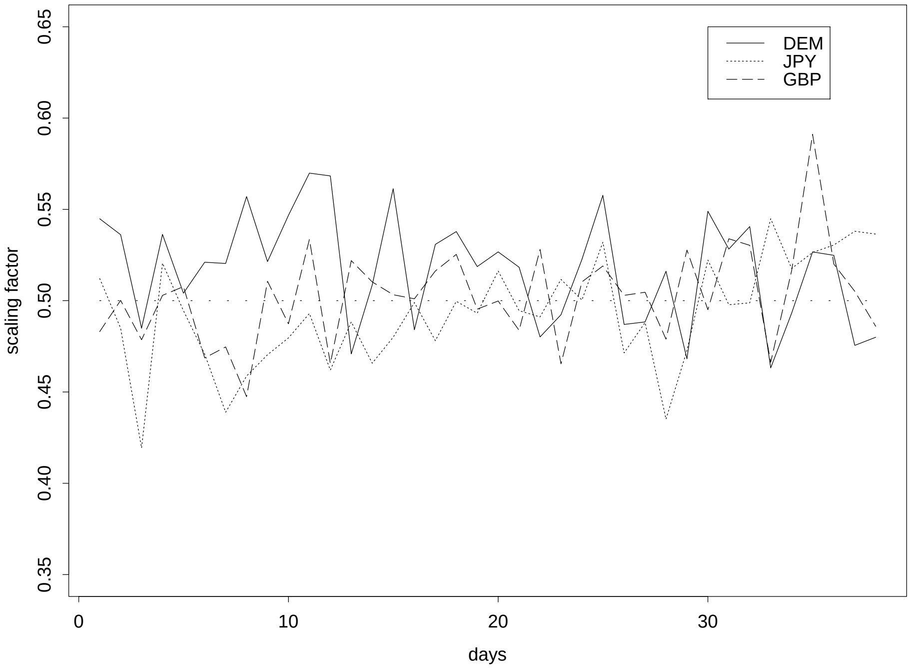

Figure 9 Empirical time aggregation rule (non-overlapping case) – power parameter plot with respect to quantile frequency for some exchange rates with $p = 0.01$. The dashed line at 0.5 corresponds to the power parameter when applying the square-root-of-time rule.

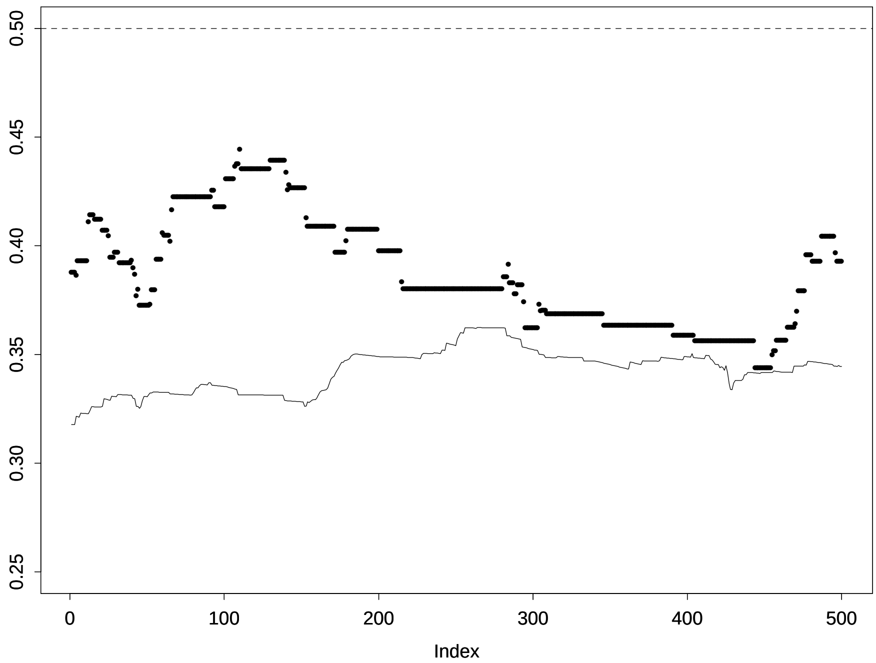

Figure 10 Hill plot (dots), empirical power parameter (line) for the DAX. The dashed line at 0.5 corresponds to the power parameter when applying the square-root-of-time rule.

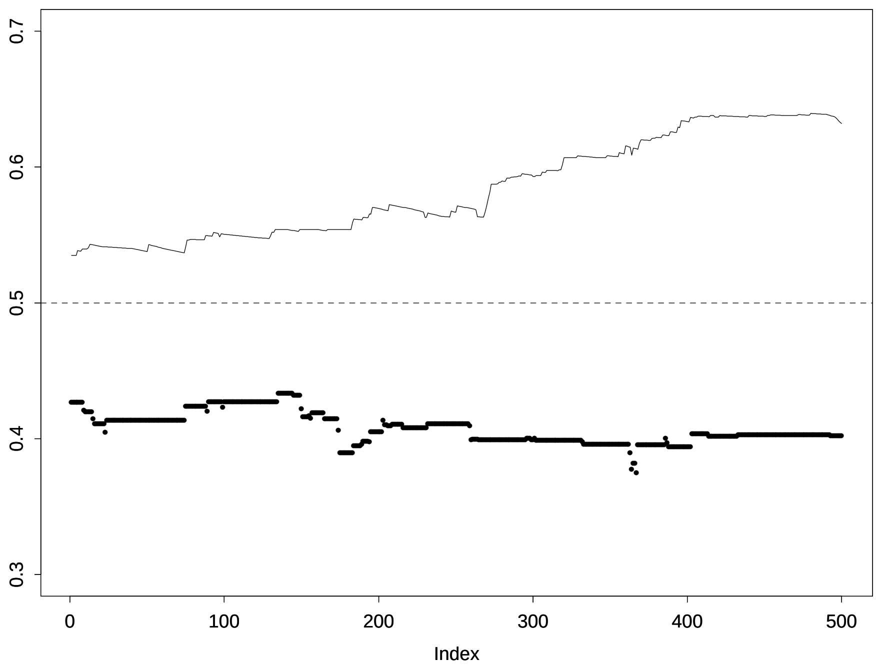

Figure 11 Hill plot (dots), empirical power parameter (line) for USD/DEM exchange rates. The dashed line at 0.5 corresponds to the power parameter when applying the square-root-of-time rule.

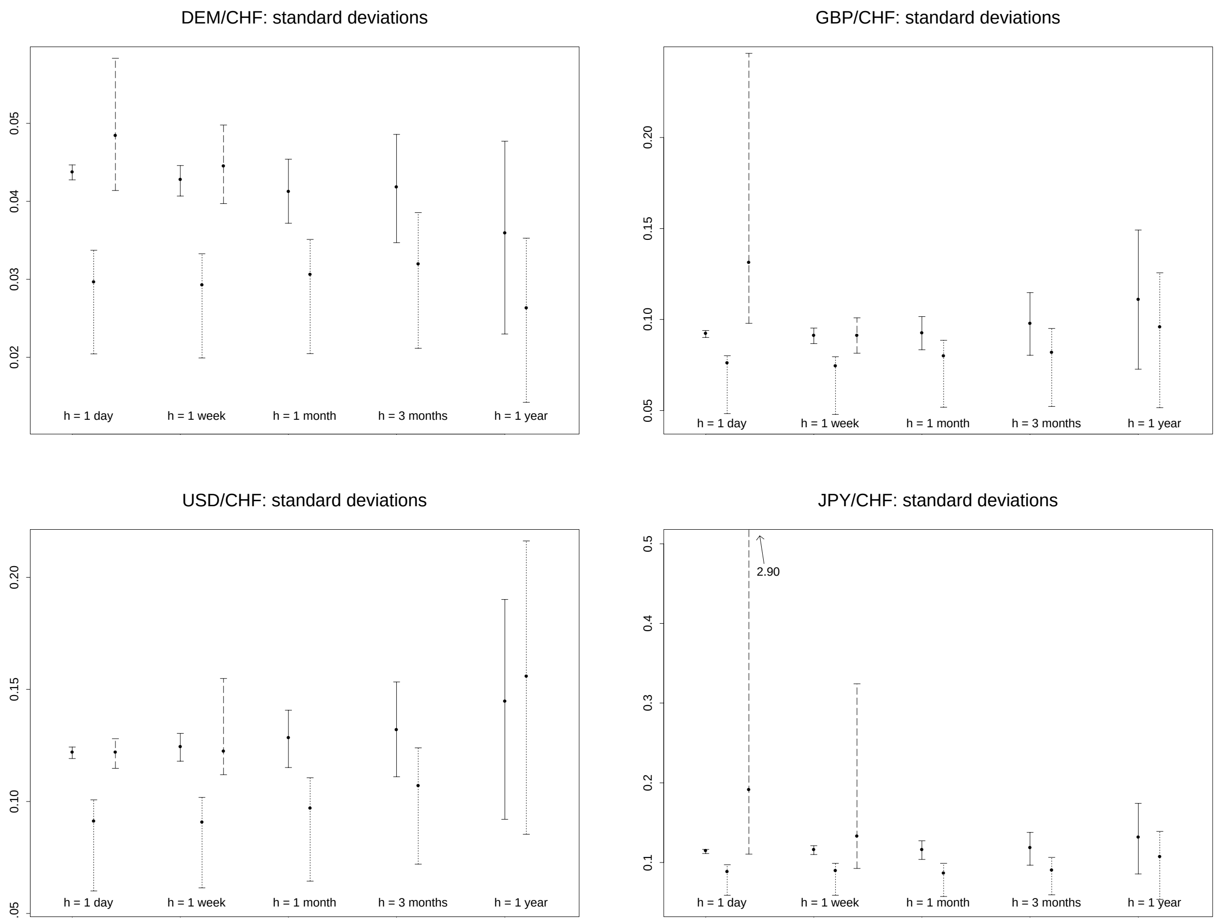

Figure 12 Foreign exchange rates (DEM/CHF, GBP/CHF, USD/CHF, JPY/CHF): estimates for standard deviations (including confidence intervals constructed via simulation), using random walk models (full lines), AR(p) models (dotted lines) and GARCH(1,1) models (dashed lines).

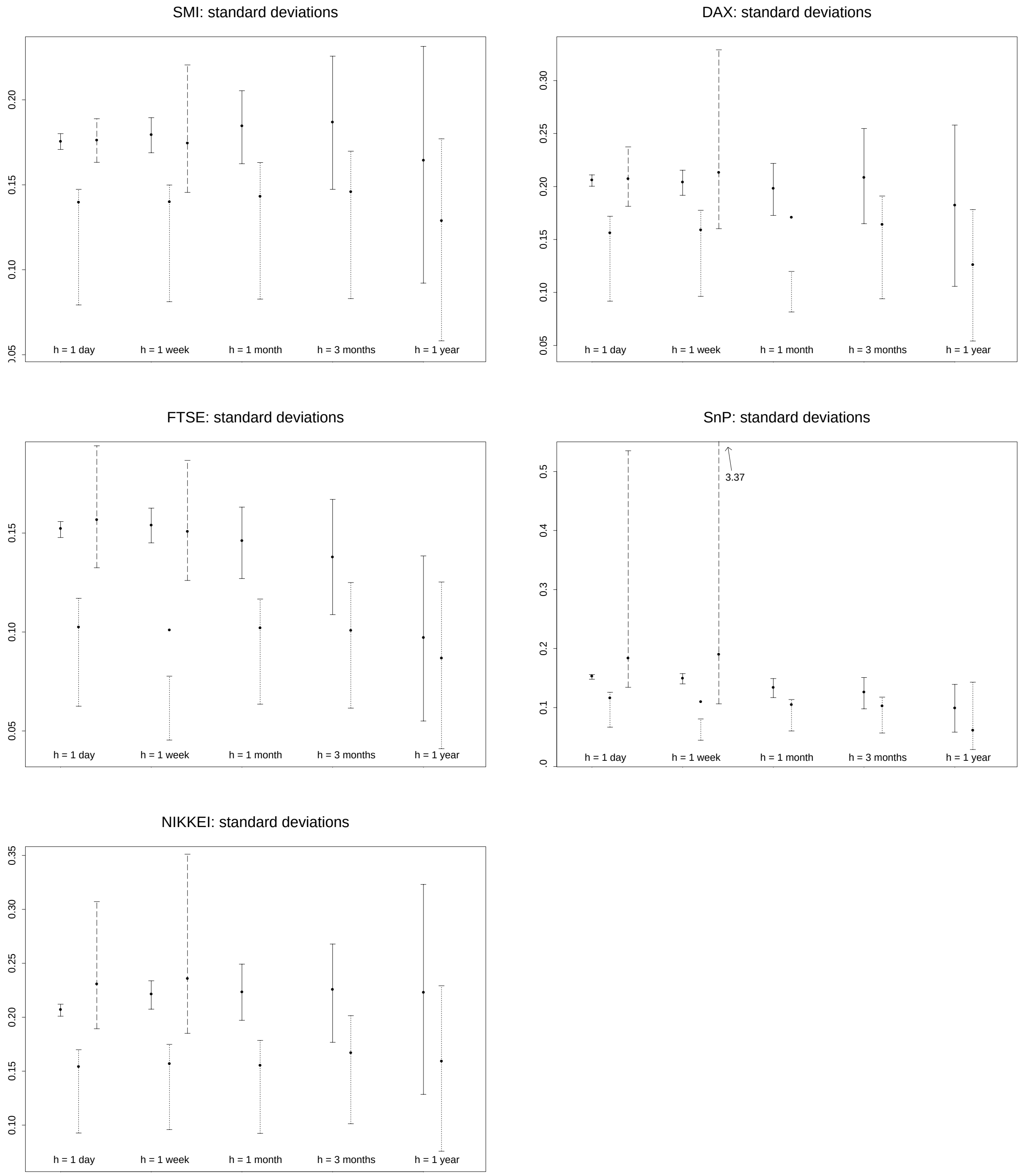

Figure 13 Stock indices (SMI, DAX, FTSE, S&P, NIKKEI): estimates for standard deviations (including confidence intervals constructed via simulation), using random walk models (full lines), AR(p) models (dotted lines) and GARCH(1,1) models (dashed lines).

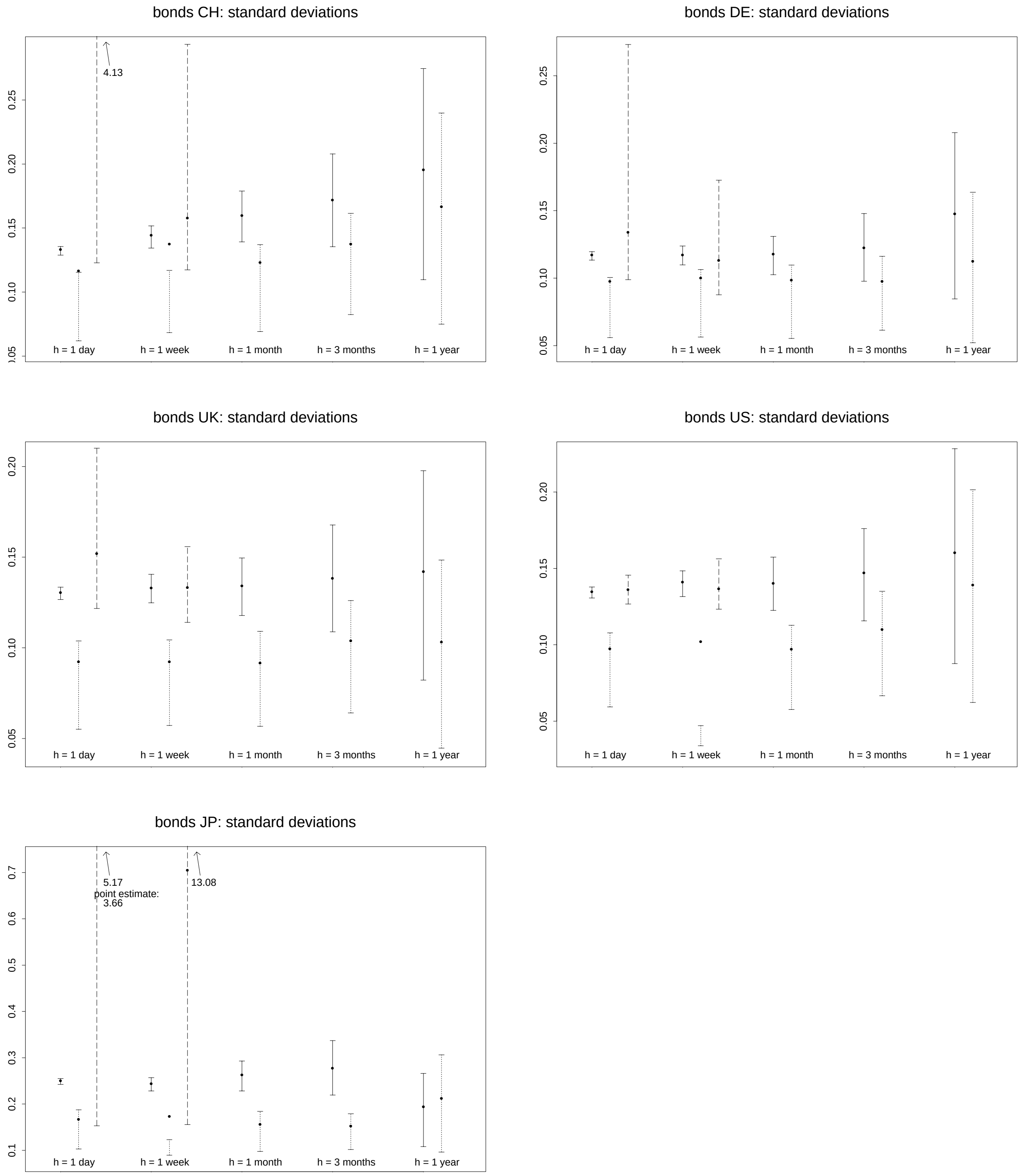

Figure 14 10-year government bonds (CH, DE, UK, US, JP): estimates for standard deviations (including confidence intervals constructed via simulation), using random walk models (full lines), AR(p) models (dotted lines) and GARCH(1,1) models (dashed lines).

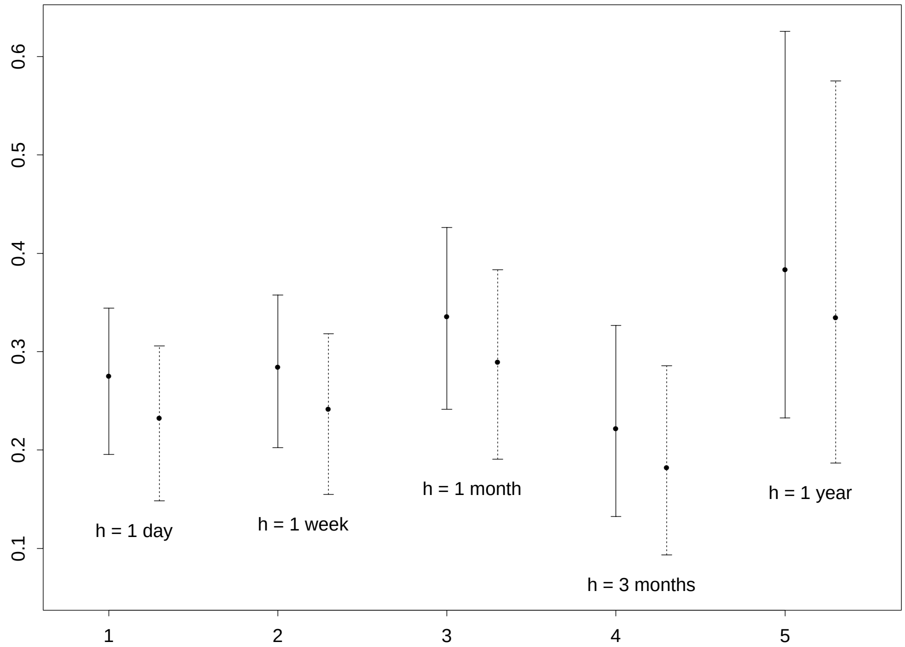

Figure 15 Point estimates and 95% confidence intervals for one-year 1% expected shortfall and 1% value-at-risk for the SMI (*percentage loss*), based on five different intermediate horizons h. Underlying model: random walk with normal innovations.

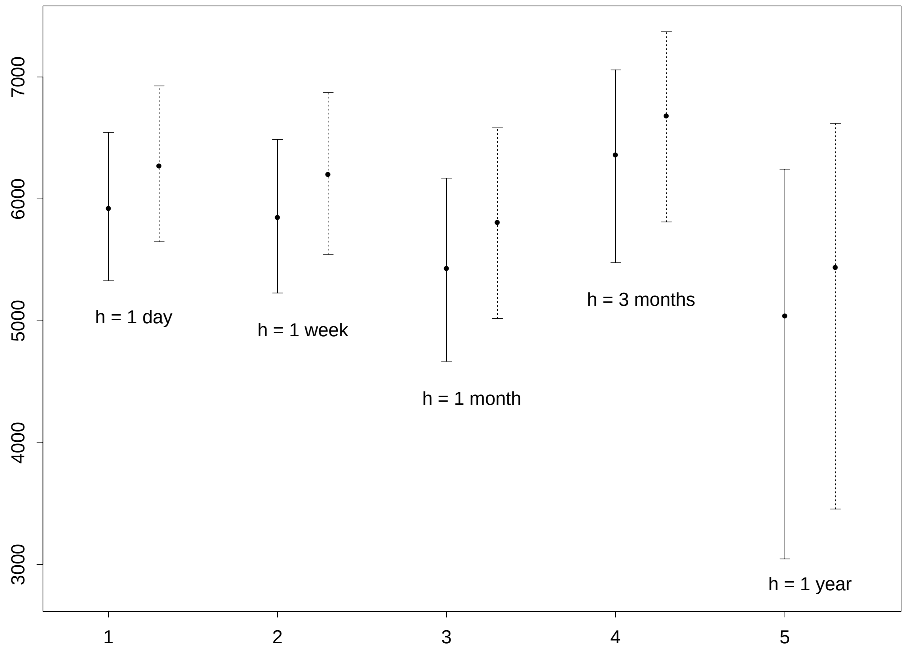

Figure 16 Point estimates and 95% confidence intervals for one-year 1% expected shortfall and 1% value-at-risk for the SMI, transformed in such a way that the values at the end of the year 2001 can be read off. The estimates are based on five different intermediate horizons h. Underlying model: random walk with normal innovations.

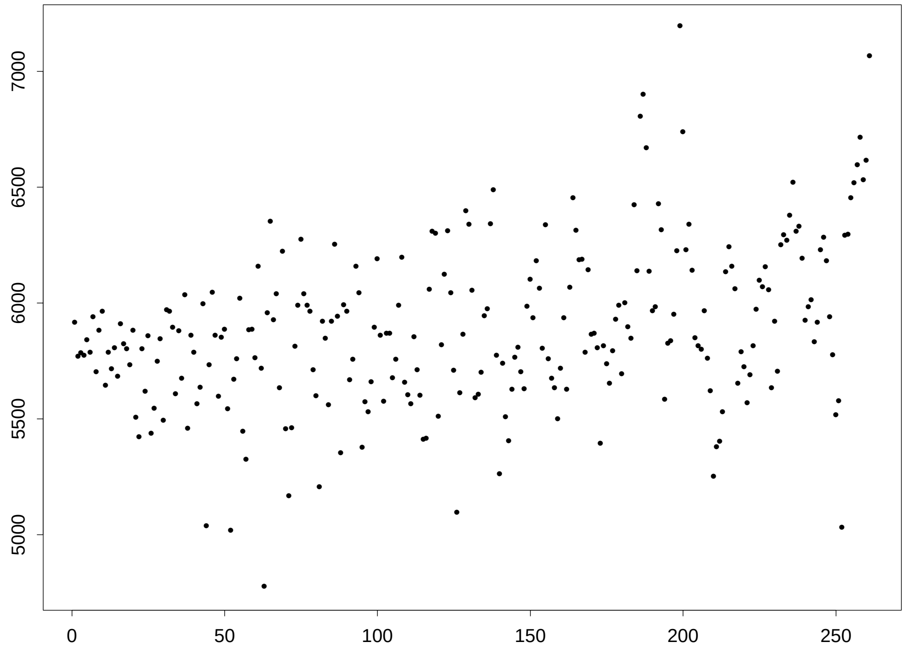

Figure 17 Point estimates for one-year 1% expected shortfall for the SMI, transformed in such a way that the values at the end of the year 2001 can be read off. The estimates are based on five intermediate horizons h from one day to one year. Underlying model: random walk with normal innovations.

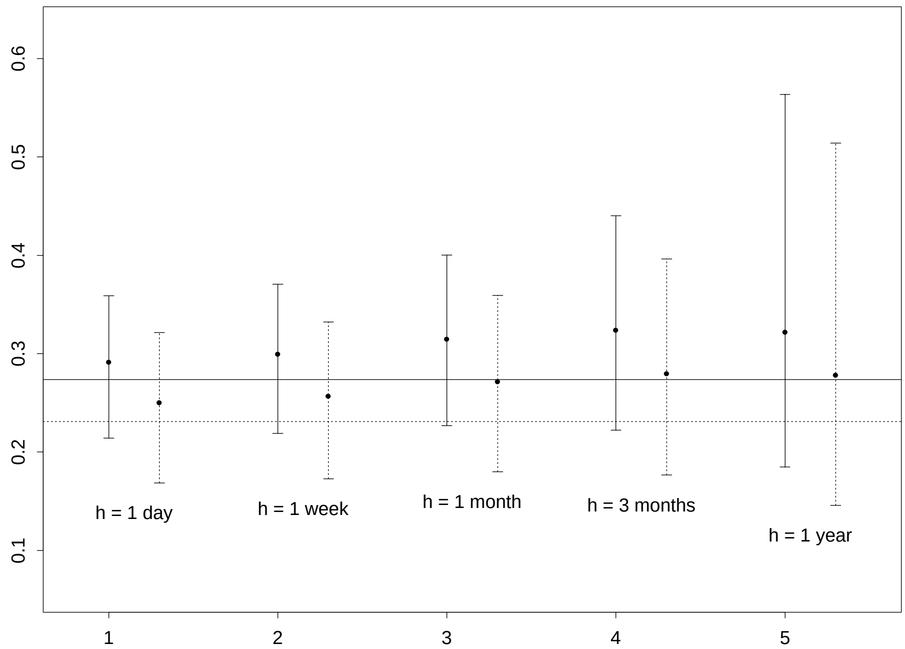

Figure 18 Point estimates and 95% confidence intervals for one-year 1% expected shortfall and 1% value-at-risk (percentage loss) for a simulated random walk, based on five different intermediate horizons h. Underlying model: random walk with normal innovations. The true values are 27.35% for expected shortfall and 23.09% for value-at-risk.

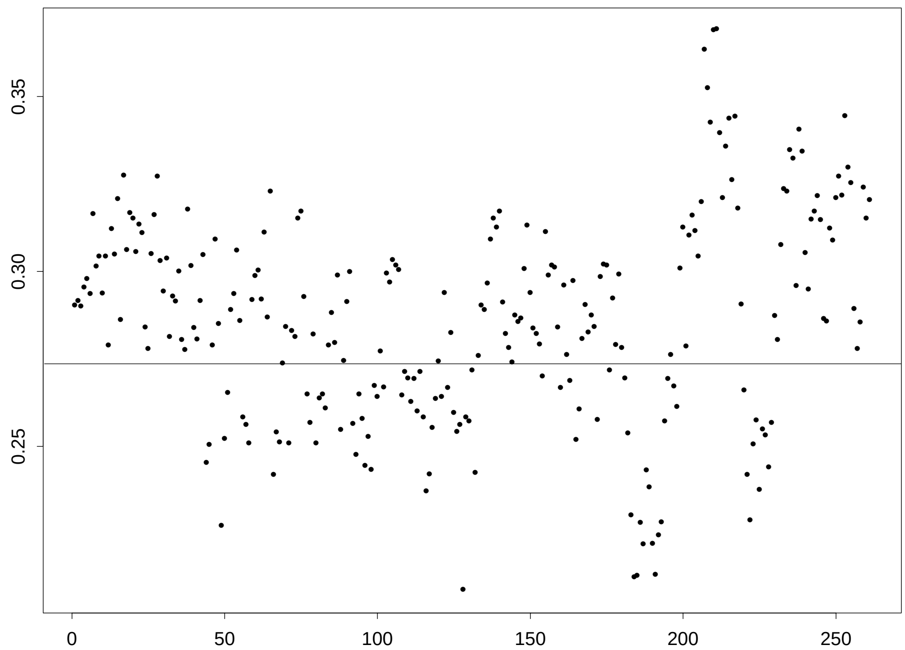

Figure 19 Point estimates for one-year 1% expected shortfall (percentage loss) for a simulated random walk, based on five intermediate horizons h from one day to one year. Underlying model: random walk with normal innovations. The true value is 27.35%.