Overview of Models Measuring Long-Term Financial Risks

Roger Kaufmann, Pierre Patie

RiskLab

ETH Zürich

June 22, 2000

http://www.risklab.ch/Papers.html#SLTFR

http://www.math.ethz.ch/∼kaufmann

http://www.math.ethz.ch/∼patie

Overview of Models Measuring Long-Term Financial Risks

R. Kaufmann, P. Patie

Part 1

Part 2

I General Description

- Scope
- Risk measure definition
- Key points

Scope

The goal is the development of a theoretically well-understood and empirically founded conceptual framework for the measurement of long-term financial risk of strategic investment portfolios.

We focus on modelling the stochastic evolution of risk factors associated to portfolio positions over a long time horizon. We then estimate the long-term (e.g. 1 month, 1 year) risk of such a portfolio using expected shortfall as our reference risk measure. We hold the portfolio weights constant.

Risk measure definition

Definition 1 *Let $\mathcal{G}$ be the set of all risks, a measure of risk is a mapping from $\mathcal{G}$ into $\mathbb{R}$.*

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition 2 *Given $\alpha \in]0,1[$, the value-at-risk VaR_α at level α of the returns R with distribution $\mathbb{P}$, is*

$$VaR_\alpha(R) = -\inf\{x \in \mathbb{R} \mid \mathbb{P}[R \le x] \ge \alpha\},$$

i.e. VaR is the negative of the α-quantile of R.

Then, we consider as risk measure the **expected shortfall**.

Definition 3 *The expected shortfall ES_α at a level α is defined by*

$$ES_\alpha = -\mathbb{E}[R \mid R < -VaR_\alpha], \text{ where } R \in \mathbb{L}^1(\mathbb{P}).$$

The expected shortfall is a *coherent* risk measure in the sense of Artzner, Delbaen, Eber and Heath.
Value-at-risk is not !

Key questions

1. Which frequency do we use to fit models?
    - Are long datasets stationary?
    - What are the statistical restrictions?
    - How can we keep as much information as possible?
2. Do the properties of financial data change when we choose another time horizon?
3. What is the reliability of the time aggregation rule of each model if there are any?
4. How can we compare different models?

II Empirical Studies

- Data exploration
- Graphical tools
- Long-range dependence or non-stationarity?
- Empirical time aggregation function

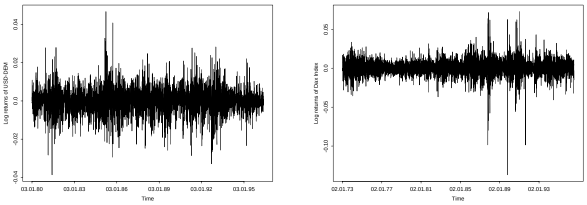
Data Exploratory

We explore:

- log-returns of USD/DEM exchange rate from 03.01.80 to 21.05.96,
- log-returns German stock index DAX (Deutscher Aktienindex) from 02.01.73 to 23.07.96.

Data	Freq.	N.	Mean	Var.	Skew.	Kurt.	Min	Max
USD/DEM	1 day	4274	−0.0001	0.00005	0.0058	5.93	−0.039	0.047
	1 week	854	−0.0005	0.0002	−0.1749	5.96	−0.087	0.075
	2 weeks	427	−0.0009	0.0005	−0.6666	6.36	−0.145	0.071
	1 month	213	−0.0019	0.001	−0.2399	4.61	−0.151	0.105
DAX	1 day	6146	0.00025	0.0001	−0.75	16.02	−0.137	0.073
	1 week	1229	0.0012	0.0006	−0.63	7.03	−0.168	0.099
	2 weeks	614	0.0025	0.001	−0.38	4.60	−0.152	0.109
	1 month	307	0.0051	0.002	−0.642	4.70	−0.199	0.140

Graphical tools

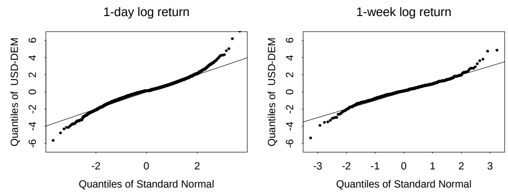


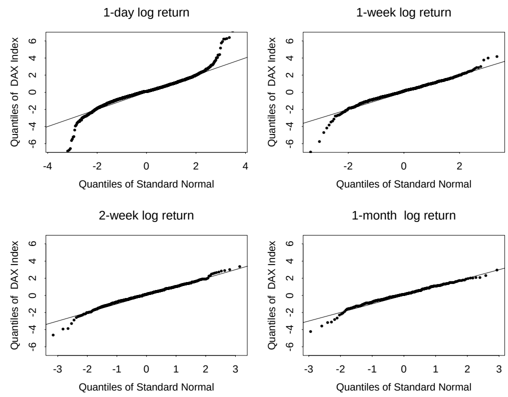
Log returns of USD / DEM exchange rate and of DAX index.

The quantile-quantile-plot (*QQ-plot*) against the normal distribution.
It is defined by:

$$\left\{ \left(N^{-1} \left(\frac{n-k+1}{n+1} \right), \tilde{r}^h_{k,n} \right), k = 1, \ldots, n \right\},$$

where $\tilde{r}^h_{k,n}$ denotes the normalized and centered kth order statistic of the h-days log-returns and N^{-1} is the inverse of the standard Gaussian distribution function.





long range dependence or non-stationarity?

Assuming that $\mathbb{E}(r_t^2) < +\infty$, we plot

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

where

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

and

$$\gamma_{n,|\mathbf{r}|}(h) = \frac{1}{n} \sum_{t=1}^{n-h} |r_t||r_{t+h}| - \bar{\mathbf{r}}^2, \qquad h \in \mathbb{N},$$

$$v_{|\mathbf{r}|}(h) = \overline{|r_0 r_h|}^2, \qquad h \in \mathbb{N}.$$

Let p_j, $j = 0, \ldots, k$, be positive numbers such that $p_1 + \ldots + p_k = 1$ and $p_0 = 0$. Define

$$q_j = p_0 + \ldots + p_j, \quad j = 0, \ldots, k.$$

Assume the sample $X_1, \ldots, X_n$ consists of k subsamples

$$X_1^{(1)}, \ldots, X_{[n\,q_1]}^{(1)}, \ldots, X_{[n\,q_k]+1}^{(k)}, \ldots, X_n^{(k)}.$$

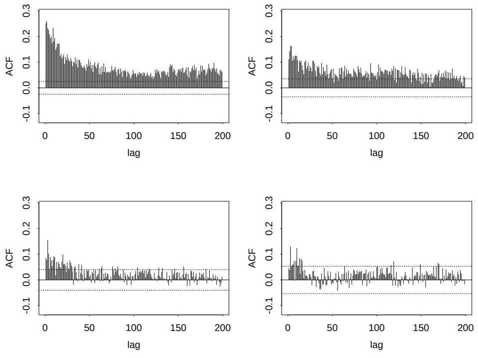
The i^{th} subsample comes from a stationary ergodic model with finite 2^{nd} moment.

By the ergodic theorem it follows that for fixed $h \geq 0$ as $n \to \infty$

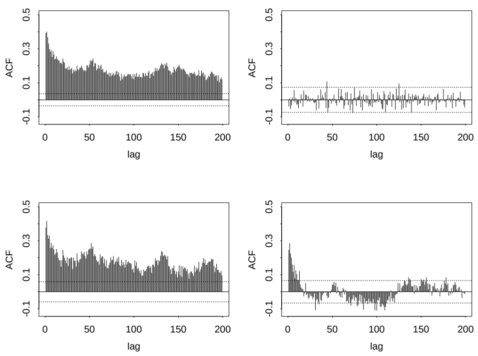
$$\begin{aligned} \tilde{\gamma}_{n,X}(h) &= \sum_{j=1}^{k} p_j \frac{1}{np_j} \sum_{t=[nq_{j-1}]+1}^{[nq_j]} X_t^{(j)} X_{t+h}^{(j)} \\ &\quad - \left(\sum_{j=1}^{k} p_j \frac{1}{np_j} \sum_{t=[nq_{j-1}]+1}^{[nq_j]} X_t^{(j)} \right)^2 + o(1) \\ &\longrightarrow \sum_{j=1}^{k} p_j \gamma_{X^{(j)}}(h) \\ &\quad + \sum_{1 \leq i \leq j \leq k} p_i p_j \left(\mathbb{E}(X^{(j)}) - \mathbb{E}(X^{(i)}) \right)^2 \text{ a.s.} \end{aligned} \qquad (1)$$

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".

An explanation of this feature may be a type of non-stationarity, i.e. shifts in the unconditional volatility of the model underlying the log-returns.



ul:whole sample; ur:73-90; ll:75-85; lr:75-81.



Empirical time aggregation function (1)

We are interested in finding an empirical relation between quantile values computed at the same level but for different frequencies.
Consider $(r_t)_{1 \leq t \leq N}$ to be a strictly stationary time series, representing N observations of the log returns on a financial asset price.
We denote by r_h^p the p-quantile value defined by

$$r_h^p := \inf\{r \in \mathbb{R} \mid \mathbb{P}[\sum_{i=1}^{h} r_i \leq r] \geq p\}.$$

Formally, we can write the quantile relation down as follows:

for $1 \leq h' \leq h$ and $p \in [0,1]$,

$$r_h^p = \left(\frac{h}{h'}\right)^{\epsilon_{h,h'}} r_{h'}^p,$$

thus,

$$\epsilon_{h,h'} = \log\left(\frac{r_h^p}{r_{h'}^p}\right)\left(\log\left(\frac{h}{h'}\right)\right)^{-1}, \qquad (2)$$

Empirical time aggregation function (2)

We first choose as reference quantile the quantile obtained from the empirical distribution of daily log-returns ($h' = 1$).

- we pick ns data up,
- we estimate quantiles for different values of h ($h = 2 \ldots 250$),
- we compute $\epsilon_{h,h'}$ by using (2).

We repeat this procedure k times ($k = N - ns - 1$).
We define

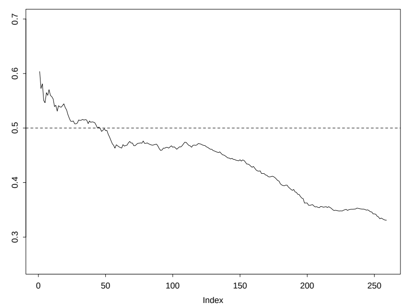
$$\bar{\epsilon}_{h,1} = \frac{1}{k}\sum_{i=1}^{k} \epsilon_{h,1}^{(i)}. \qquad (3)$$

Then we plot for each dataset

$$\left\{(h, \bar{\epsilon}_{h,1}),\ h = 2 \ldots 250\right\}.$$

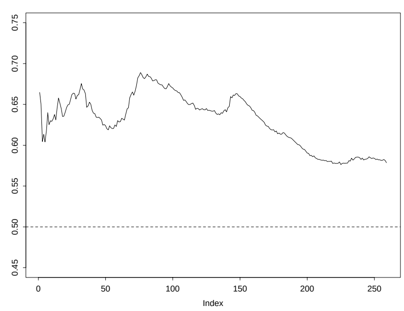
We notice that for DAX index:

- for $h \leq 10$ days, $\epsilon_{h,1} > 0.5$,
- for $h > 3$ months, $0.3 < \epsilon_{h,1} < 0.4$.



Power parameter for Dax index (line),

$\epsilon = 0.5$ (dash)



Power parameter for USD/DEM (line),

$\epsilon = 0.5$ (dash)

III Model Description

- Generalized hyperbolic distribution
- Heavy-tailed distribution
- Vector auto regressive model with error correction
- Stochastic volatility models
- GARCH
- Random Walk

Generalized hyperbolic distributions

GHD were introduced by Barndorff-Nielsen (1977).

$$h_{\lambda,\alpha,\beta,\delta,\mu}(x) = a(\lambda,\alpha,\delta,\mu)\left(\delta^2 + (x-\mu)^2\right)^{(\lambda-\frac{1}{2})/2} K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)\exp\left(\beta(x-\mu)\right), \quad (4)$$

where

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

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

These densities depend on five parameters:

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

Generalized inverse Gaussian distributions

A special case of GHD (when $\lambda = -\frac{1}{2}$) is the normal inverse Gaussian distribution.
It was introduced to finance by Barndorff-Nielsen.

$$g_{\lambda,\alpha,\beta,\delta,\mu}(x) = \frac{\alpha}{\pi}\exp\left(\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)\right)\frac{K_1\left(\alpha\delta\sqrt{1+\left(\frac{x-\mu}{\delta}\right)^2}\right)}{\sqrt{1+\left(\frac{x-\mu}{\delta}\right)^2}}. \quad (5)$$

This class of distributions is the only sub-class to get the following convolution property:

$$g(\alpha,\beta,\delta_1,\mu_1) * g(\alpha,\beta,\delta_2,\mu_2) = g(\alpha,\beta,\delta_1+\delta_2,\mu_1+\mu_2). \quad (6)$$

In this sense it is close to the normal distribution where means and variances of independent random variables add up too.
Eberlein give analytical properties and in particular the moment generating function which allow to derive different moments.

Heavy-tailed distributions

We consider $(r_t, t \in \mathbb{N})$ to be a strictly stationary time series, independent and identically distributed (i.i.d.), representing observations of the log returns on a financial asset price.
Further, we assume

$$\mathbb{P}[r_1 < -x] = Cx^{-\alpha}L(x) \quad \text{as } x \to \infty, \quad (7)$$

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

$$\forall t > 0 : \lim_{x\to\infty}\frac{L(tx)}{L(x)} = 1.$$

These distributions are called *heavy-tailed distributions*, since the kth moment is unbounded for $\alpha < k$.
(7) is a characterisation of the maximum domain of attraction of the Fréchet distribution.

Using conditional maximum likelihood estimate we find the **Hill estimator** (*Hill, 1975*)

$$\widehat{\alpha}_{k,n} = \left(\frac{1}{k} \sum_{i=1}^{k} \log \left(\frac{r_{i,n}}{r_{k,n}} \right) \right)^{-1}, \qquad (8)$$

and

$$\widehat{C}_{k,n} = \frac{k}{n} r_{k,n}, \qquad (9)$$

where n is the sample size and $r_{k,n}$ is the kth order statistics, taken as a threshold.
Under a second order regular variation condition on L, $\widehat{\alpha}_{k,n}$ is asymptotically normally distributed with asymptotic variance α^2.

By inverting (7) we can easily derive unconditional measure estimates for the p-quantile

$$\widehat{x}^p_{k,n} = r_{k,n} \left(\frac{kp}{n} \right)^{\frac{1}{\widehat{\alpha}_{k,n}}}, \qquad (10)$$

and for the corresponding expected shortfall

$$\widehat{ES}^p_{k,n} = \frac{\widehat{\alpha}_{k,n} + 1}{\widehat{\alpha}_{k,n} - 1} \widehat{x}^p_{k,n}, \quad \widehat{\alpha}_{k,n} > 1. \qquad (11)$$

Time aggregation

Feller's theorem (1971)

Theorem 4 *We assume that* $(r_t, t \in \mathbb{N})$ *have a distribution as in* (7)*.*
As $x \to \infty$*, we have*

$$\mathbb{P}\left[\sum_{i=1}^{h} r_i < -x \right] = hCx^{-\alpha}[1 + o(1)], \qquad (12)$$

where the scale factor C *is as in* (7)*.*

$\sum_{i=1}^{h} r_i$ can be considered as the h-days log returns.
When applicable, this theorem supplements the central limit theorem by providing information concerning the tails.

Dacorogna, Müller, Pictet and de Vries present the following rule:

Theorem 5 *Suppose* r *has a finite variance (i.e.* $\alpha > 2$*). At a constant risk level* p*, increasing the time horizon* h *increases the VaR and the expected shortfall numbers for the heavy tailed model by a factor* $h^{\frac{1}{\alpha}}$*.*

From now, we assume that the following second order expansion applies: as $x \to \infty$

$$\mathbb{P}[r_1 < -x] = Cx^{-\alpha}[1 + bx^{-\beta} + o(1)], \quad C, b > 0, \quad \alpha, \beta > 0. \qquad (13)$$

From this second order expansion, de Vries derives asymptotic properties of the Hill estimator.

Lemma 6 *The asymptotic bias of the Hill estimator* (8) *is*

$$\mathbb{E}\left[\frac{1}{\widehat{\alpha}_{k,n}} - \frac{1}{\alpha} \right] = -\frac{b\beta}{\alpha(\alpha + \beta)} r_{k,n}^{-\beta} + o(r_{k,n}^{-\beta}).$$

The asymptotic variance of the Hill estimator for the threshold $r_{k,n} \to \infty$*, and* $\frac{r^{\alpha}_{k,n}}{n} \to 0$*,*

$$Var\left[\frac{1}{\widehat{\alpha}_{k,n}} - \frac{1}{\alpha} \right] = \frac{r^{\alpha}_{k,n}}{C\, n\, \alpha^2} + o\left(\frac{r^{\alpha}_{k,n}}{n} \right).$$

We obtain the asymptotic mean squared error (AMSE)

$$AMSE\left(\frac{1}{\widehat{\alpha}_{k,n}} \right) \approx \frac{r^{\alpha}_{k,n}}{C\, n\, \alpha^2} + \frac{b^2 \beta^2}{\alpha^2 (\alpha + \beta)^2} r_{k,n}^{-2\beta}. \qquad (14)$$

Proposition 7 *Suppose the* r_i *are i.i.d. with a distribution function* $F(x)$ *that is symmetric around zero, and is as in* (13) *with* $\alpha > 2$*. Then a* h*-convolution affects the leading term in the AMSE* (14) *of the Hill estimator as follows:*

1. $\beta < 2$*. There is no effect.*

2. $\beta = 2$*. The AMSE changes by a factor*

$$\left[\left(1 + \frac{1}{2} \alpha(\alpha + 1)(h - 1)\mathbb{E}[r_1^2]/b \right)^2 \right]^{\alpha/(2\beta+\alpha)}.$$

3. $\beta > 2$*. The AMSE changes by a factor*

$$\rho \left[\frac{1}{2} \alpha(\alpha + 1)(h - 1)\mathbb{E}[r_1^2] \right]^{\frac{2\alpha}{4+\alpha}} \left(\frac{1}{b^2} \right)^{\alpha/(2\beta+\alpha)},$$

where

$$\rho = \frac{4 + \alpha}{2\beta + \alpha} \left(\frac{2}{\alpha + 2} \right)^{\frac{2\alpha}{4+\alpha}} \left(\frac{\alpha + \beta}{\beta} \right)^{\frac{2\alpha}{2\beta+\alpha}} \left(\frac{\alpha}{4\, C\, n} \right)^{\frac{4}{4+\alpha}} \left(\frac{2\beta\, C\, n}{\alpha} \right)^{\frac{2\beta}{2\beta+\alpha}}.$$

Test for EVT time aggregation rule

Let's recall the time aggregation rule for heavy-tailed distribution:

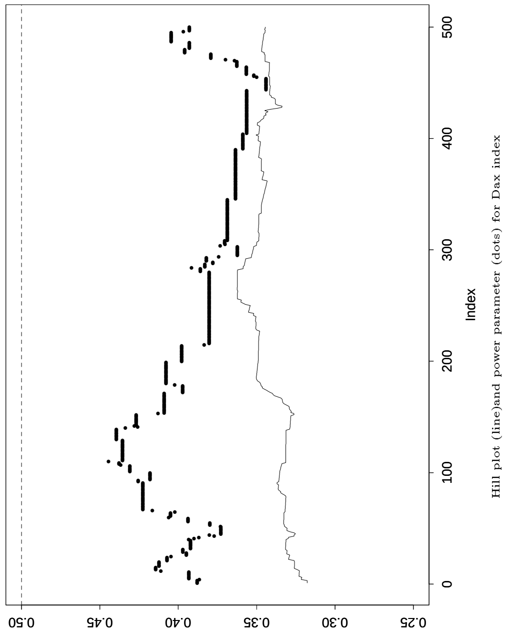
Theorem 8 *Suppose r has a finite variance (i.e. $\alpha > 2$). At a constant risk level p, increasing the time horizon h increases the VaR and the expected shortfall numbers for the heavy tailed model by a factor $h^{\frac{1}{\alpha}}$.*

We compare for the same frequency ($h = 250$) the empirical power parameter and the Hill estimator. We plot

$$\left\{(k, \epsilon^k_{250,1}), k = 1 \dots 500\right\},$$

and

$$\left\{(k, \widehat{\alpha}^k), k = 1 \dots 500\right\}$$



Forecasts Based on Economic Structure

Problem: Nonparametric models require (too) large amounts of data. Therefore we have to restrict on parametric models.

VECM = VARM (Vector Auto Regressive Model)
+ ECM (Error Correction Model)

VARM with p lags:

$$\mathbf{s}_t = \Theta\,\mathbf{d}_t + \Gamma_1\,\mathbf{s}_{t-1} + \cdots + \Gamma_p\,\mathbf{s}_{t-p} + \epsilon_t,$$

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,
- $\epsilon_t \overset{\text{iid}}{\sim} \mathcal{N}_n(0, \Lambda)$,
- Λ : covariance matrix.

Error Correction Model (ECM)

Idea: Existence of stable relationships between variables.

These relationships provide an error correction mechanism.

ECM for n nonstationary variables and k cointegration relations:

$$\mathbf{s}_t = A\,\mathbf{v}_t + \boldsymbol{\epsilon}^{\mathbf{s}}_t,$$
$$\mathbf{v}_t = \mathbf{v}_{t-1} + \boldsymbol{\epsilon}^{\mathbf{v}}_t,$$

where

- $\mathbf{s}_t, \boldsymbol{\epsilon}^{\mathbf{s}}_t$: n-dimensional stochastic vectors,
- $\mathbf{v}_t, \boldsymbol{\epsilon}^{\mathbf{v}}_t$: $(n-k)$-dimensional stochastic vectors,
- A : $n \times (n-k)$ matrix,
- $\boldsymbol{\epsilon}^{\mathbf{s}}_t, \boldsymbol{\epsilon}^{\mathbf{v}}_t$: uncorrelated stationary random variables.

We can rewrite the ECM in the form

$$\mathbf{s}_t = \mathbf{s}_{t-1} - \boldsymbol{\epsilon}^{\mathbf{s}}_{t-1} + \boldsymbol{\epsilon}^{\mathbf{s}}_t + A\,(\mathbf{v}_t - \mathbf{v}_{t-1}),$$
$$\mathbf{v}_t = \mathbf{v}_{t-1} + \boldsymbol{\epsilon}^{\mathbf{v}}_t.$$

Error Correction Model: An Example

With

$$\mathbf{s}_t := \begin{pmatrix} s_{1,t} \\ s_{2,t} \end{pmatrix},$$
$$A := \begin{pmatrix} a \\ 1 \end{pmatrix},$$
$$\boldsymbol{\epsilon}_t^{\mathbf{s}} := \begin{pmatrix} \epsilon_{1,t}^{\mathbf{s}} \\ 0 \end{pmatrix},$$
$$\mathbf{v}_t := s_{2,t},$$
$$\epsilon_{1,t}^{\mathbf{s}}, \epsilon_t^{\mathbf{v}} \overset{\text{iid}}{\sim} \mathcal{N}(0,1),$$

i.e.

$$s_{1,t} = a\, s_{2,t} + \epsilon_{1,t}^{\mathbf{s}},$$
$$s_{2,t} = s_{2,t-1} + \epsilon_t^{\mathbf{v}},$$

we get

$$s_{1,t} = s_{1,t-1} - \epsilon_{1,t-1}^{\mathbf{s}} + \epsilon_{1,t}^{\mathbf{s}} + a\,(s_{2,t} - s_{2,t-1}),$$
$$s_{2,t} = s_{2,t-1} + \epsilon_t^{\mathbf{v}}.$$

Hence for positive values of a an increase in $s_{2,t}$ leads to a higher value of $s_{1,t}$.

Vector Auto Regressive Model with Error Correction (VECM)

The VECM is a combination of two models:

- Vector Auto Regressive Model (VARM)
$$\mathbf{s}_t = \Theta\,\mathbf{d}_t + \Gamma_1\,\mathbf{s}_{t-1} + \cdots + \Gamma_p\,\mathbf{s}_{t-p} + \boldsymbol{\epsilon}_t.$$

- Error Correction Model (ECM)
$$\mathbf{s}_t = A\,\mathbf{v}_t + \boldsymbol{\epsilon}_t^{\mathbf{s}},$$
$$\mathbf{v}_t = \mathbf{v}_{t-1} + \boldsymbol{\epsilon}_t^{\mathbf{v}}.$$

VECM Formula

$$\Delta\mathbf{s}_t = \Theta\,\mathbf{d}_t + \Pi\,\mathbf{s}_{t-1} + \sum_{j=1}^{p-1} \Gamma_j\,\Delta\mathbf{s}_{t-j} + \boldsymbol{\epsilon}_t,$$

where

$\Delta\mathbf{s}_t = \mathbf{s}_t - \mathbf{s}_{t-1}$
n-dimensional stochastic vector of log returns,
$\mathbf{d}_t$: vector of deterministic variables,
Θ : matrix of coefficients,
Γ_j : $n \times n$ matrix,
$\Pi = \alpha\,\beta^T$: $n \times n$ matrix,
α, β : $n \times k$ matrices,
$\boldsymbol{\epsilon}_t \overset{\text{iid}}{\sim} \mathcal{N}_n(0, \Lambda)$,
Λ : covariance matrix.

The structure of VARM is obvious. The ECM part is included in the term $\Pi\,\mathbf{s}_{t-1}$. Since the matrix Π can be decomposed into $\Pi = \alpha\,\beta^T$, we can write $\Pi\,\mathbf{s}_{t-1}$ in the form $\alpha\,\beta^T\mathbf{s}_{t-1}$, where $(\beta^T\mathbf{s}_{t-1})$ are the transformed stationary series.

Scaling Rule: Square Root of Time

Let s_t be the log price at time t. We assume that changes in the log price are independently and identically distributed,

$$s_t - s_{t-1} = \epsilon_t,$$

where

ϵ_t are iid, with $\mathbb{E}[\epsilon_t] = 0$, $\mathbb{E}[\epsilon_t^2] = \sigma^2$.

Then the h-day log returns can be written as

$$s_t - s_{t-h} = \sum_{i=0}^{h-1} \epsilon_{t-i},$$

which has mean zero and variance $h\sigma^2$. Hence the "square root of time" rule works to convert 1-day volatility to h-day volatility in the i.i.d. case.

In practice we are more interested in converting quantile estimates. Hence for applying the square root of time rule, we furthermore have to require that the log returns are normally distributed.

Random Walk Models

- RWZ: random walk with zero expected log return.
- RWF: random walk with expected log return equal to the forward premium.
- RWF_{imp}: random walk with expected log return equal to the forward premium, and using option implied volatility.

Random Walk with Zero Expected Log Return (RWZ)

- $S_t =$ asset price $(t \in \mathbb{N}_0)$
- $s_t = \log S_t$

Assumption 1: Expected log returns are equal to zero

$$\mathbb{E}_t[s_{t+1}] - s_t = 0.$$

Assumption 2: Normally distributed, independent log returns with standard deviation σ in each period $[t, t+1]$.

$$\text{Ass. 1 \& 2} \Rightarrow s_{t+1} - s_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2).$$

→ The logarithmic asset price follows a random walk with zero drift.

σ^2 can be estimated by

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{i=0}^{T-1}(s_{t-i} - s_{t-i-1})^2, \qquad T \le t.$$

Question:

- Modelling of interest rate?

Random Walk with Expected Log Return Equal to the Forward Premium (RWF)

Assumption 1': Expected log returns are equal to the forward premium

$$\mathbb{E}_t[s_{t+1}] - s_t = r_{t,t+1} - d_{t,t+1},$$

where

$r_{t,t+1} =$ risk free continuously compounded interest rate for period $[t, t+1]$,

and $d_{t,t+1}$ is defined by

$S_{t+1}(e^{d_{t,t+1}} - 1) =$ asset's 1-period cash payouts at time $t+1$, known at t.

Assumption 1' can be rewritten as follows:

$$\mathbb{E}_t[\log(S_{t+1}\, e^{d_{t,t+1}}\, e^{-r_{t,t+1}})] = \log(S_t).$$

Assumption 2: Normally distributed, independent log returns with standard deviation σ in each period $[t, t+1]$.

$$\text{Ass. 1' \& 2} \Rightarrow s_{t+1} - s_t \sim \mathcal{N}(r_{t,t+1} - d_{t,t+1}, \sigma^2).$$

→ The logarithmic asset price follows a random walk with drift equal to the forward premium. The drift for $[t, t+1]$ is random and time-varying, but known at time t.

Estimation of σ in the RWF Model

If forward price data is available, σ^2 can be estimated by

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{i=0}^{T-1}(s_{t-i} - f_{t-i-1,t-i})^2, \qquad T \le t,$$

where $f_{t,t+1}$ is the *logarithmic forward price*:

$$\begin{aligned} f_{t,t+1} &= \log(S_t\, e^{r_{t,t+1} - d_{t,t+1}}) \\ &= s_t + r_{t,t+1} - d_{t,t+1}. \end{aligned}$$

If forward price data is not available, σ^2 can be estimated by

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{i=0}^{T-1}(s_{t-i} - s_{t-i-1})^2, \qquad T \le t.$$

But since we neglect the drift in this formula, we get a biased estimator.

Random Walk with Expected Log Return Equal to the Forward Premium, and Using Option Implied Volatility (RWF$_{\text{imp}}$)

Instead of using historical volatility we can also use *option implied volatility* to forecast the variance.

There is no explicit formula for calculating option implied volatilities, but iterative procedures based on the Black-Scholes formula can be used.

Value of a Whole Portfolio

distribution of individual prices and rates

⇓ add dependence structure

complete multivariate model

⇓ simulation of price paths

monthly values of the whole portfolio

⇓ use Brownian Bridge process

daily values of the whole portfolio, i.e. information about the distribution of the portfolio change in value

Stochastic Volatility Models

- $S_t =$ asset price.
- $X_t = \log\left(\frac{S_t}{S_{t-1}}\right) - \mu$
  centered daily log returns.

Stochastic volatility:

$$X_t = \sigma_t\,\epsilon_t, \quad \mathbb{E}[\epsilon_t] = 0,\ \mathbb{E}[\epsilon_t^2] = 1,\ (\epsilon_t) \text{ iid.}$$

- Markovian stochastic volatility models:

$$\sigma_t = F(\sigma_{t-1}, Z_t, t),$$
$$Z_t = \text{white noise (i.e. iid), independent of } (\epsilon_t)_{t\in\mathbb{N}}.$$

- Generalization:

$$\sigma_t = F(\sigma_{t-1}, X_{t-1}, Z_t, t).$$

Stochastic Volatility Models (2)

- Regime-switching volatility:

  $$\sigma_t = v_{u(t)}, \quad u(t)\text{: state.}$$

  Transition probabilities given by matrix Π

  $$\Pi_{u_1,u_2} = \mathbb{P}[u(t+1) = u_2 \,|\, u(t) = u_1].$$

- Log-auto-regressive volatility

  $$\log\sigma_t^2 = \alpha + \gamma\,\log\sigma_{t-1}^2 + \kappa\,Z_t,$$

  $Z_t \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$,
  α, γ, κ constants, $\gamma \in (-1,1)$.

- GARCH(1,1)

  $$\sigma_t^2 = \alpha_0 + \alpha_1\,X_{t-1}^2 + \beta_1\,\sigma_{t-1}^2,$$

  $\alpha_0, \alpha_1, \beta_1 > 0$,
  stationarity condition: $\alpha_1 + \beta_1 < 1$.

Stochastic Volatility Models (3)

- EGARCH

$$\log \sigma_t^2 = \alpha + \gamma \log \sigma_{t-1}^2 + \beta_1 \frac{X_{t-1}}{\sigma_{t-1}} + \beta_2 \left(\left| \frac{X_{t-1}}{\sigma_{t-1}} \right| - \sqrt{\frac{2}{\pi}} \right).$$

- Cross-market GARCH

2-market-version:

$$\begin{pmatrix} \sigma_{a,t}^2 \\ \sigma_{ab,t} \\ \sigma_{b,t}^2 \end{pmatrix} = \alpha + \beta \begin{pmatrix} X_{a,t-1}^2 \\ X_{a,t-1} X_{b,t-1} \\ X_{b,t-1}^2 \end{pmatrix} + \gamma \begin{pmatrix} \sigma_{a,t-1}^2 \\ \sigma_{ab,t-1} \\ \sigma_{b,t-1}^2 \end{pmatrix},$$

$X_{a,t}$: centered log return in market a at time t,
$X_{b,t}$: centered log return in market b at time t,
$\sigma_{a,t}$: conditional volatility of $X_{a,t}$,
$\sigma_{b,t}$: conditional volatility of $X_{b,t}$,
$\sigma_{ab,t}$: conditional covariance between $X_{a,t}$ and $X_{b,t}$,
β, γ: 3×3 matrices.

Stochastic Volatility Models (4):

- *Jump-Diffusion Models*

Plain vanilla:

$$X_t = \sigma_t \, \epsilon_t,$$

where

σ_t : deterministic or stochastic,
$\epsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$.

Discretized jump-diffusion:

$$X_t = \sigma_t \, \epsilon_t + N_t \, \tilde{\sigma}_t \, \tilde{\epsilon}_t,$$

where

$\sigma_t, \tilde{\sigma}_t$: deterministic or stochastic,
$\epsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$,
$\tilde{\epsilon}_t$ iid, independent of $(\epsilon_t)_{t \in \mathbb{N}}$,
$N_t \overset{\text{iid}}{\sim} Be(p)$.

GARCH(1,1)

Let $(X_t, t \in \mathbb{N})$ be a strictly stationary time series representing observations of centered log returns on a financial asset price.

A GARCH(1,1) model for X is defined by

$$X_t = \sigma_t \, \epsilon_t \quad \text{for } t \in \mathbb{N},$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$
$$\epsilon_t \text{ iid}, \ \mathbb{E}[\epsilon_t] = 0, \ \mathbb{E}[\epsilon_t^2] = 1.$$

Stationarity conditions:
$0 < \alpha_0 < \infty$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$ and $\alpha_1 + \beta_1 < 1$.

Fit the GARCH(1,1) process by pseudo-maximum-likelihood estimation to obtain the value of the parameters of the conditional volatility.

Scaling Rule: GARCH Coefficients Function

Assume: Centered 1-day log returns X_t follow a GARCH(1,1) process with a normal distributed random term.

$$X_t = \sigma_t \, \epsilon_t,$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$
$$\epsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0,1).$$

Drost-Nijman:

$$X_{h,t} := \sum_{i=0}^{h-1} X_{t-i} \quad \text{is GARCH(1,1):}$$

$$X_{h,t} = \sigma_{h,t} \, \epsilon_{h,t},$$
$$\sigma_{h,t}^2 = \alpha_{h,0} + \alpha_{h,1} X_{h,t-1}^2 + \beta_{h,1} \sigma_{h,t-1}^2,$$
$$\epsilon_{h,t} \overset{\text{iid}}{\sim} \mathcal{N}(0,1),$$
$$\alpha_{h,1} \to 0, \beta_{h,1} \to 0 \quad \text{as } h \to \infty.$$

Test of the Drost-Nijman Formula

We examine whether the Drost-Nijman formula gives good approximations for real data.

Parameter estimation for a GARCH(1,1) model with a normal distributed random term:

- Deutscher Aktienindex (DAX)

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

- USD / DEM exchange rate

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

There exist two possibilities for estimating the GARCH(1,1) parameters of h-day returns:

1. Use the formula by Drost and Nijman.
2. Estimate these parameters directly from data.

Deutscher Aktienindex (DAX)		
$h = 5$ days (1 week)	Drost-Nijman	$\sigma_{5,t}^2 = 6.586 \cdot 10^{-5} + 0.10485\, X_{5,t-1}^2 + 0.7924\, \sigma_{5,t-1}^2$
	direct estimation	$\sigma_{5,t}^2 = 1.751 \cdot 10^{-5} + 0.09731\, X_{5,t-1}^2 + 0.8710\, \sigma_{5,t-1}^2$
$h = 20$ days (1 month)	Drost-Nijman	$\sigma_{20,t}^2 = 9.023 \cdot 10^{-4} + 0.09640\, X_{20,t-1}^2 + 0.5519\, \sigma_{20,t-1}^2$
	direct estimation	$\sigma_{20,t}^2 = 5.740 \cdot 10^{-5} + 0.09217\, X_{20,t-1}^2 + 0.8847\, \sigma_{20,t-1}^2$
$h = 80$ days (4 months)	Drost-Nijman	$\sigma_{80,t}^2 = 8.449 \cdot 10^{-3} + 0.04016\, X_{80,t-1}^2 + 0.1364\, \sigma_{80,t-1}^2$
	direct estimation	no result
$h = 261$ days (1 year)	Drost-Nijman	$\sigma_{261,t}^2 = 3.336 \cdot 10^{-2} + 0.00665\, X_{261,t-1}^2 - 0.0032\, \sigma_{261,t-1}^2$
	direct estimation	no result
$h \to \infty$	Drost-Nijman	$\sigma_{h,t}^2 = 1.283 \cdot 10^{-4}\, h$

USD / DEM exchange rate		
$h = 5$ days (1 week)	Drost-Nijman	$\sigma_{5,t}^2 = 1.097 \cdot 10^{-5} + 0.06977\, X_{5,t-1}^2 + 0.8840\, \sigma_{5,t-1}^2$
	direct estimation	$\sigma_{5,t}^2 = 1.035 \cdot 10^{-5} + 0.08140\, X_{5,t-1}^2 + 0.8785\, \sigma_{5,t-1}^2$
$h = 20$ days (1 month)	Drost-Nijman	$\sigma_{20,t}^2 = 1.637 \cdot 10^{-4} + 0.08110\, X_{20,t-1}^2 + 0.7463\, \sigma_{20,t-1}^2$
	direct estimation	$\sigma_{20,t}^2 = 1.419 \cdot 10^{-4} + 0.16233\, X_{20,t-1}^2 + 0.7152\, \sigma_{20,t-1}^2$
$h = 80$ days (4 months)	Drost-Nijman	$\sigma_{80,t}^2 = 2.016 \cdot 10^{-3} + 0.05766\, X_{80,t-1}^2 + 0.4109\, \sigma_{80,t-1}^2$
	direct estimation	no result
$h = 261$ days (1 year)	Drost-Nijman	$\sigma_{261,t}^2 = 1.133 \cdot 10^{-2} + 0.01835\, X_{261,t-1}^2 + 0.0660\, \sigma_{261,t-1}^2$
	direct estimation	no result
$h \to \infty$	Drost-Nijman	$\sigma_{h,t}^2 = 4.742 \cdot 10^{-5}\, h$

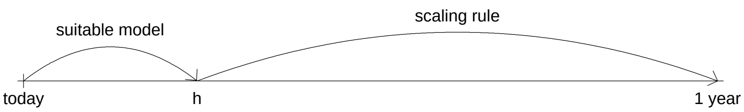
IV Model Comparison

- Test methodology
- Backtesting description
- Results

Model Comparison

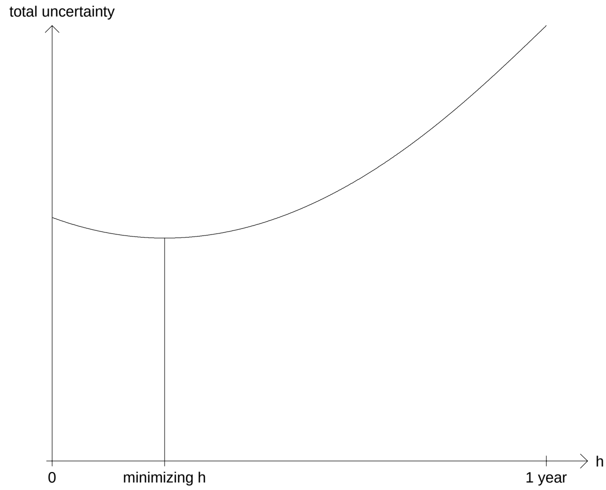
No one of the proposed models obviously outperforms the others. Each of them has its deficiencies.

All models only perform well for relatively short time horizons. We have to fix a horizon $h <$ 1 year, for which we can use our models. For the gap between h and 1 year, we will have to use a scaling rule.



Optimal Time Horizon

Modelling h-day log returns leads to a first uncertainty. The scaling of 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 order to find then the best model among the ones we investigate in, we have to compare their minimal total uncertainties.

Backtesting: Test Description

- *Frequency of exceedence*
  It consists of calculating the percentage of times that the observed returns fall below the negative VaR estimate:

$$V^{\text{freq}} = \frac{1}{t_1 - t_0 + 1} \sum_{t=t_0}^{t_1} 1_{\{R_t < -\text{VaR}_t\}}.$$

- *Relative amount of exceedence*
  We count the relative excess beyond the VaR focusing on the relative size of the exceeding real values:

$$V^{\text{size}} = \frac{\sum_{t=t_0}^{t_1} \frac{R_t - (-\text{VaR}_t)}{-\text{VaR}_t} 1_{\{R_t < -\text{VaR}_t\}}}{\sum_{t=t_0}^{t_1} 1_{\{R_t < -\text{VaR}_t\}}}.$$

- Finally, we can combine these two measures by creating a *weighted average indicator*

$$V = \lambda V^{\text{freq}} + (1 - \lambda) V^{\text{size}} \quad \text{for a } \lambda \in [0, 1].$$

Comparing Time Horizons in a GARCH(1,1) Model

For different time horizons, we compare

- difference between empirical expected shortfall and estimated expected shortfall,
- frequency of real data exceeding estimated value-at-risk,
- relative amount of exceedences.

For this comparison, we look at prices of 10 german stocks from 02.01.73 to 23.07.96
(i.e. 6146 daily log returns).

Horizon	Bias of ES	Frequency	Relative Amount of Exceedences
1 day	0.0731	2.918%	0.1916
2 days	0.0736	2.760%	0.2007
5 days	0.0797	2.784%	0.2165
10 days	0.0800	2.740%	0.2189
20 days	0.0779	2.591%	0.2196

V Conclusions

- The models described reflect the behaviour seen in real financial data.
- Some of the models are interesting in the sense that they have some theoretical time aggregation rule. Thus they can be fitted to relatively high frequency data. On the one hand this allows to keep as much information as possible and on the other hand it avoids the problem of (non)stationarity.
- Indeed it is not evident that financial data are stationary over a long time period.
- The behaviour of exchange rate data and stock data differ at some points (empirical time aggregation function).
- Next steps:
  - fit the models,
  - compare them by using our test methodology.

References

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

Barndorff-Nielsen, O.E. (1998) *Processes of normal inverse Gaussian type*, Finance & Stochastics, 2:41-68 (1998).

Christoffersen P.F., Diebold F.X. and Schuermann T. (1998) *Horizon Problems and Extreme Events in Financial Risk Management*, Financial Institutions Center 98-16, The Wharton School, University of Pennsylvania.

Dacorogna M.M., Müller U.A., Pictet O.V. and de Vries C.G. (1999), *Extremal Forex Returns in Extremely Large Data Sets*.

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

Eberlein E. (1999) *Application of generalized hyperbolic Lévy motions to finance*, Universität Freiburg, FDM preprint.

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

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

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

RiskMetrics (1999) *LongRun Technical Document*, available from the web:
http://www.riskmetrics.com/research/techdocs/splash.html#Corp