Gürtler, Marc; Rauh, Ronald

Working Paper

Shortcomings of a parametric VaR approach and nonparametric improvements based on a non-stationary return series model

Working papers // Institut für Finanzwirtschaft, Technische Universität Braunschweig, No. IF32V2

Provided in cooperation with:
Technische Universität Braunschweig

Working Paper Series

Shortcomings of a Parametric VaR Approach and Nonparametric Improvements Based on a Non-stationary Return Series Model

by Marc Gürtler and Ronald Rauh

No.: IF32V2/09
First Draft: 2009-05-18
This Version: 2009-09-09

Technische Universität Braunschweig
Institute of Finance
Abt-Jerusalem-Str. 7
D-38106 Braunschweig

Shortcomings of a Parametric VaR Approach and Nonparametric Improvements Based on a Non-stationary Return Series Model

by Marc Gürtler* and Ronald Rauh**

Abstract. In this paper we first investigate the validity of a general Value at Risk approach, which is widely used for risk management in banking and insurance companies. We discuss and widely reject the conventional assumptions, e.g. independent identically distributed normal returns, and as consequence develop an improved model for non-stationary returns. Therein volatility dynamics are modelled both exogenously and deterministic, captured by a nonparametric regression-type approach. Consistency and asymptotic normality of a symmetric and of a one-sided kernel estimator of volatility are outlined with remarks on the bandwidth decision. We pay further attention to asymmetry and heavy tails of the return distribution, implemented by the framework for innovations. On a multitude of financial time series for equity indices, exchange rates, interest rates and credit spreads it is shown that the univariate approach is practically manageable and outperforms the standard tools.

Keywords: delta-normal model, volatility, Value at Risk (VaR), heteroscedastic asset returns, non-stationarity, nonparametric regression, innovation modelling, forecasting, backtesting

JEL classification: C12, C14, C15, C5

---

* **Professor Dr. Marc Gürtler**
Technische Universität Braunschweig
Institute of Finance
Abt-Jerusalem-Str. 7
D-38106 Braunschweig
Germany
Fon: +49 531 3912895
Fax: +49 531 3912899
E-mail: marc.guertler@tu-bs.de

** **Dipl.-Math. Oec. Ronald Rauh**
Öffentliche Versicherung Braunschweig
Department of Asset Risk Management
Theodor-Heuss-Str. 10
D-38122 Braunschweig
Germany
Fon: +49 531 2021568
Fax: +49 531 202331568
E-mail: ronald.rauh@oeffentliche.de

1 Introduction

Financial risk management has become an important task for banks and insurance companies in the last two decades. Focussing on the European market, first the Basel Committee on Banking Supervision defined the Value at Risk (henceforth VaR) as the benchmark risk measure for a banks capital requirements. Similar concepts are currently developed by the Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS) for the insurance sector. But only the smaller part of VaR calculations are made for regulatory purposes, it is widely used for estimating the riskiness of trading activities, for the maximum loss potential in market value of an existing portfolio and for minimization of risk in portfolio optimization. A growing importance of risk management is observed in an asset-liability framework, focussing on a total market value balance and the surplus between assets and liabilities. Especially in life insurance, it is a main target to keep the duration gap between fixed income securities and long-term liabilities with an obligatory target yield low or rather in line with a risk budget to avoid reinvestment problems.

In this paper we first investigate the validity of a general Value at Risk approach, which is widely used for risk management in financial institutions. We discuss and widely reject the conventional assumptions, e.g. independent identically distributed normal returns, and as consequence develop an improved model for non-stationary returns. Therein volatility dynamics are modelled both exogenously and deterministic, captured by a nonparametric regression-type approach. Consistency and asymptotic normality of a symmetric and of a one-sided kernel estimator of volatility are outlined with remarks on the bandwidth decision. We pay further attention to asymmetry and heavy tails of the return distribution, implemented by the framework for innovations. On a multitude of financial time series for equity indices, exchange rates, interest rates and credit spreads it is shown that our univariate approach is practically manageable and outperforms the standard tools.

In the early days of time series analysis, a return series $\{X_t\}_{t=1,\ldots,n}$ on market prices $\{P_t\}_{t=0,\ldots,n}$ of an asset over n periods of time was expected to be a series of independent and identically distributed (iid) random variables and $\ln P_t = \ln P_{t-1} + X_t$ was assumed to follow a random walk; an overview gives Samuelson (1973). He proposed modelling prices in continuous time by a geometric Brownian motion, whose discretization leads to a random walk with iid normal increments for the discrete log-prices. The normal distribution assumption was preserved in a plenty of practical applications due to its great features (e.g. invariance property) until modern days of finance: For example, the Markowitz portfolio theory (Markowitz (1952)) postulates multivariate normal yields. Black and Scholes (1973), or Merton (1973) respectively, built up their option pricing theory on the geometric Brownian motion of prices. Moreover, we will start later on with a basic risk model, that exploits the mathematical tractability of iid normal returns.

The hypothesis of geometric Brownian motion was first rejected in the 1960s statistically, e.g. by the analyses of Mandelbrot (1963) and Fama (1965) on several US stock series. They draw the following empirical conclusions on return series, often labeled as *stylized facts*:

- there is a serial data dependence,
- the volatility is changing over time (heteroscedasticity),
- the returns are asymmetrically distributed with heavy tails,
- a negative return amplitude entails a greater volatility than a positive return of same amount (Leverage Effects).

Concerning the last stylized fact, Nelson (1991) gives theoretical considerations that volatility tends to response asymmetrically to positive and negative returns. We later provide evidence on the stylized facts as a by-product of testing the assumptions of the basic risk model on a multitude of financial time series. An indication for the first stylized fact is often derived from correlograms, that test graphically whether there is any sample autocorrelation $\rho(h) = \gamma(h)/\gamma(0)$, with $\gamma(h)$ the empirical estimate for $Cov\,(X_t, X_{t+h})$, for several discrete time lags $h \geq 0$, provided a stationary series. The typical picture for liquid financial instruments is that daily returns themselves contain little serial correlation, but their absolute returns $|X_t|$ are significantly positive correlated over a large number of lags with a slowly declining sample autocorrelation function (SACF). For stock returns Taylor (1986) established that there is little serial correlation, which is in agreement with the efficient market theory, and that absolute and squared returns contain substantially more correlation.[1] This effect is also named *long range dependence* (LRD), and one might draw the conclusion of a serial dependence of the time series.

[1] Ding et al. (1993) extended the result to power transformations of absolute returns, $|X_t|^d$ for $d > 0$, and showed on an S&P return series them to be 'long-memory', with quite high autocorrelations for large lags. Furthermore, they found that for a fixed lag τ the function $\rho_\tau(d) = Corr\left(|X_t|^d, |X_{t+\tau}|^d\right)$ has a unique maximum when $d \approx 1$.

On the ideas of serial dependence and conditional time-varying volatility Robert F. Engle and Tim Bollerslev developed in the 1980s nonlinear time series models as *autoregressive conditional heteroscedastic* (ARCH-processes). The general form is

$$X_t = \mu_t + \varsigma_t \varepsilon_t, \quad t \in \mathbb{Z}, \tag{1}$$

where $\{\varepsilon_t\}_t$ is an iid sequence with $E\varepsilon_1 = 0$ and $Var\varepsilon_1 = 1$, $\{\mu_t\}$ and $\{\varsigma_t\}$ are stochastic processes that depend only on past information, i.e. $\mu_{t+1}, \varsigma_{t+1}$ are measurable with respect to the σ-field $\mathcal{F}_t = \sigma(\{\varepsilon_j \mid j \leq t\})$. Hence X_t is $\mathcal{F}_t$-measurable for every t and $\mu_{t+1} = E(X_{t+1} \mid \mathcal{F}_t)$ is the conditional mean and $\varsigma_{t+1}^2 = Var(X_{t+1} \mid \mathcal{F}_t)$ is the conditional variance of X_{t+1} given the past returns. The starting point was set by the seminal paper of Engle (1982) introducing the ARCH(p)-process, where conditional volatility dynamics $\{\varsigma_t\}_t$ are imposed via linear regression over past squared (centered) returns. An extension of Bollerslev (1986), named generalized ARCH (GARCH(p,q)-process), includes past variances next to historical returns into the parametric regressed volatility process. This enables often a parsimonious parametrization with a reasonable fit to empirical data. For purposes of parameter estimation, processes of the ARCH-family are defined to be stationary. But finding conditions for the existence and uniqueness of a stationary solution was nontrivial.[2] Besides the focus on heteroscedasticity and uncorrelated, serial dependence, the features of heavy-tailedness and asymmetry are mostly imposed on the distribution of innovations ε_t. The exponential GARCH (EGARCH) model of Nelson (1991) and the asymmetric power GARCH (AGARCH) model by Ding et al. (1993) include asymmetry and leverage effects directly in the volatility dynamics, that reacts different to financial gains and losses. In his nobel price lecture Engle (2004) gives a brief history of ARCH models and an alphabet of model extensions, but the 'GARCH(1,1) specification is the workhorse of financial applications' (Engle (2004), p. 408), emphasising that this is a good starting point for an analysis of multifaceted financial returns.

Coming back to the LRD, Mikosch and Starica (2004) derive theoretically that the aforementioned SACF for absolute returns could alternatively arise from non-stationarities in the data. Correlograms are only a significant tool for detecting dependence under the assumption of stationarity, otherwise structural breaks in the data as shifts in the variance might cause identical results,[3] and it cannot be discriminated between stationary long-memory and non-stationary time series. Secondly, Granger and Starica (2005) came in a long-term case study on the S&P 500 index to the conclusion, that the second interpretation is the more probable alternative: The main reason for the sample LRD is to be seen in the non-stationarities due to structural breaks of the unconditioned variance. From this follows a grave criticism on ARCH-type models, since they are parameterized as stationary processes (that implies a fortiori a fixed unconditional variance) and focus on modelling a long-range dependence structure of the second moments. More inconsistencies arise on ARCH modelling over longer periods of daily returns: So, the typical outcome of a GARCH$(1,1)$ implementation is that the sum of estimated parameters is approximately one, leading to an IGARCH$(1,1)$ model, which is referred to as *IGARCH effect* in Starica (2003) or Mikosch and Starica (2004). But an IGARCH-model implies an infinite variance of the observed random variables, which contradicts to the results of a direct tail analysis indicating that daily returns have a finite second moment (see De Haan et al. (1994)). Mikosch and Starica (2004) prove theoretically and empirically that the IGARCH effect may be generated by non-stationarities via shifts in the unconditional variance of the return series.[4] Although some ARCH-models that allow structural breaks in the volatility while holding up stationarity were developed,[5] one should question the stationarity assumption at all.

In face of that shortcomings, a huge list of ARCH-type specifications was imposed to the literature (see Engle (2004) or Bollerslev et al. (1994) for a statistical overview), where more and more sophisticated processes were developed to specify the volatility dynamics for fitting the return data. Following Drees and Starica (2002), the need of an increasing complexity for volatility modelling can be possibly explained that a simple endogenous specification does not exist. In that case, the model fit can only be improved by a change of the working hypothesis: in their univariate approach the volatility is supposed to be exogenous to the return process. The evolution of market prices is interpreted as a manifestation of complex market conditions, driven by unknown exogenous factors. In the multiplicative model

[2] Bougerol and Picard (1992) gave the solution for stationary GARCH(p,q) processes via stochastic recurrence equations. Straumann (2004) summarizes this among other stochastic features on GARCH processes.

[3] The authors found that the stronger the non-stationarity, e.g. the difference of the variation of subsamples $X^{(1)}$ and $X^{(2)}$ measured as $\left(E|X^{(1)}| - E|X^{(2)}|\right)^2$, the more pronounced the LRD effect is. This theoretical result is supplemented with a long-term empirical study of the S&P 500 index by ex-/ including the 1970s US-recession, generating the LRD effect.

[4] Concerning persistence in variance and long memory (IGARCH- and LRD effect) caused by structural changes, see also Lamoureux and Lastrapes (1990) and Diebold and Inoue (2001).

[5] E.g. regime-switching ARCH models by Hamilton and Susmel (1994) with transitions governed by an unobserved, fixed Markov chain.

(1) the volatility term is replaced by an unconditional variable. The corresponding variance $\sigma(t)^2$ is modelled as a smooth, deterministic function of time via a nonparametric kernel regression[6] over centered, squared returns R_t:

$$\hat{\sigma}^2(t) = \frac{\sum_{i=1}^{n} K_h(i-t) R_i^2}{\sum_{i=1}^{n} K_h(i-t)}, \quad t = 1, \ldots, n. \tag{2}$$

That approach preserves the independence assumption of log-returns, but it abandons the hypothesis of stationary, identically distributed (normal) returns. A special focus is set on an accurate description of the innovations $\{\varepsilon_t\}_t$ by fitting a Pearson type VII distribution to positive and negative innovations separately to allow asymmetry and heavy-tailedness. Drees and Starica (2002) show on a twelve-year S&P 500 example that their non-stationary model fits the data adequately and gets better short-term forecasts on the return distribution than conventional GARCH models. Herzel et al. (2005) extend these ideas to a multivariate, non-stationary framework. Vectors of financial returns are assumed to have a time-varying unconditional covariance matrix that evolves smoothly, captured by classical nonparametric regression. Vectors of standardized innovations are modelled again parametrically with Pearson VII for asymmetric, heavy tails. Another univariate extension with a time-varying expected return was developed by Mikosch and Starica (2004a).

In this article, we will support the non-stationary paradigm of Herzel et al. (2005) and Drees and Starica (2002) for modelling financial time series. The aims and scope of our research article are the following: From the starting point of the aforementioned, opposite basic risk model, which is in detail a widespread parametric VaR approach, we will motivate the necessity of an advanced non-stationary model. Since the basic risk model assumes iid normal return series, in a conception where exposure classes are mapped to benchmark indices for the purpose of VaR calculation, we apply an extensive hypothesis testing on a multitude of financial time series to reject the conventional assumptions and the random walk hypothesis, respectively. Moreover, we recognize the stylized facts and provide evidence of non-stationary and heteroscedastic returns that way. We ask for optimal return types $X_t = f(P_t, P_{t-1})$,[7] and the influence of sampling frequency on the financial return modelling.

Within the non-stationary model the main focus is set on the univariate return dynamics of financial time series, which is established by nonparametric curve estimation of the unconditional variances $\left\{\sigma^2(t)\right\}_t$ and the distributional features of the innovation series $\{\varepsilon_t\}_t$. While a symmetric volatility estimator $\hat{\sigma}(t)$ is employed to describe heteroscedasticity in the historical sample, a one-sided, historical version $\hat{\sigma}_{(1)}(t)$, that includes only past and present data, is employed for forecasting exercises. The statistical properties of nonparametric regression are executed for the symmetric case in Herzel et al. (2005) and Mikosch and Starica (2004a) referring on the results of Müller and Stadtmüller (1987). Gürtler et al. (2009) provide self-contained full proofs for both the two-sided and one-sided kernel estimators, we outline the consistency and asymptotic normality results here. Remarks on the kernel decision and an automatised bandwidth selection rule, based on cross-validation, are further subjects of interest. Following our main references, we fit the estimated innovations with an asymmetric version of the Pearson type VII distribution, to enable a flexible return modelling under aspects of asymmetry and heavy tails. But in contrast, we develop a method of moments for the parameter estimation of the one-sided Pearson VII densities.[8] By dint of a Student-t connection to the Pearson VII distribution we develop a factor-based VaR implementation of the non-stationary model, that offers a direct comparison to the basic risk model in terms of VaR forecasts. Altogether, we hope that those steps encourage a broad practical application of the non-stationary regression model.

A main focus of our research is devoted to the last objective. In simulation studies we challenge the asymptotical results on finite samples of a predefined volatility function and calibrate the complete non-stationary model to a high-volatile, heavy-tailed price process. The goodness of fit is tested in a simulation forecasting experiment and its performance is compared to the parametric VaR approach via Kupiec backtesting. In terms of extensive empirical studies on equity indices, exchange rates, interest rates and credit spreads we survey practical opportunities and drawbacks of the non-stationary regression model. We extend the existing literature by new asset types and feasible return

[6] The rescaled kernel function K_h is defined as $K_h(\cdot) = \frac{1}{h} K\left(\frac{\cdot}{h}\right)$ for an appropriate kernel K on a compact support $[-1, 1]$.

[7] In the following we use three return conceptions, corresponding to the specific models and exposure classes. The parametric risk model works, by default, with *arithmetic (or discrete) returns* $f(x, y) := (x - y)/y$ for the total returns of equity indices and the performance of exchange rates. Parallel, for interest rates and credit spreads the difference of prices between time units is regarded, called *diff-returns* $f(x, y) := x - y$. On the other hand the non-stationary return model is introduced with *log-returns* (geometric rate of return) $f(x, y) := \ln x - \ln y$ for equities and currencies, but still diff-returns for swap curve- and spread changes. In the econometric literature log-returns are prevalent, because of some nice properties (see. e.g. Jorion (2006)). The difference between arithmetic and log-returns is rather small, $X_t = \ln P_t - \ln P_{t-1} = \ln\left(1 + \frac{P_t - P_{t-1}}{P_{t-1}}\right) \approx \frac{P_t - P_{t-1}}{P_{t-1}}$, as can be seen writing $\ln(1 + x), x > -1$ as a Taylor series close to 0, especially for returns with an absolute value lower 10%.

[8] The above quoted literature uses maximum-likelihood estimation techniques.

types. In fact we recognize that the non-stationary regression modelling is problematic to be automatised at all, since the cross-validation method in nonparametric curve estimation may fail on the one hand for reasons documented in the sequel. On the other hand a trade-off between volatility estimation and Pearson VII fitting of innovations becomes obvious. We describe how the smoothing parameter and the Pearson VII parametrization may vary in time, which is to be reconsidered for forecasting purposes. Finally, we present on the example of the MSCI North America equity index an 'out-of-sample' forecasting evaluation and proof the superiority of the non-stationary paradigm against the basic risk model in terms of VaR-forecasts.

The rest of the paper is organized as follows: In section 2 we introduce a standard approach to a risk adjusted portfolio management, including a parametric VaR model. Hypothesis testing of the conventional assumptions reveals the pitfalls of that approach and motivates a non-stationary, heteroscedastic setup. We statistically introduce the multivariate regression model in section 3 and specialize on the univariate frame. Besides the statistical discussion of the two- and a one-sided kernel estimator, attention is paid to asymmetry and heavy tails in the return distribution by approximating the stochastic innovations as the Pearson type VII. We survey in section 4 whether the non-stationary regression model offers an adequate approximation to hypothetical price processes and whether it is an enhanced, practically manageable approach for the dynamics of real financial time series. We conclude in section 5.

2 A parametric VaR approach

The VaR is an estimate for the maximum loss at a predefined confidence level $(1-\alpha)$ for holding a position over a target horizon, generally,

$$VaR^{1-\alpha} = u_\alpha = \inf\{u : F(u) \geq \alpha\} \tag{3}$$

for a given positive value α close to zero, that equals to the shortfall probability. According to Jorion (2006), chapter 10, there are two main approaches to VaR measurement. *Full valuation methods* measure risk by repricing a portfolio over a multitude of scenarios, e.g. the *historical simulation method* and the *monte carlo simulation method* are well-known instances. The other group consists of *local-valuation methods*, where a portfolio is valued once at a starting point and possible movements are inferred by local derivatives. The easiest approach is the *delta-normal method*, that works with a first order Taylor approximation for the market value, e.g. the basis point value for a fixed income portfolio,[9] and infers the measure from the assumption of normally distributed risk factors. In the common literature this method is also named as *parametric Value at Risk* or *variance-covariance model*. For large portfolios with a limited insertion of options and other nonlinear financial derivatives this method provides a fast and efficient measurement of VaR. In the following we present a risk management approach, based on a parametric VaR, that is widely used in banks and insurances.

2.1 Basic risk model

An asset risk controller should possess instruments that measure the risks from a portfolio or asset strategy and compare the maximum loss amount with the (total or disposable) amount of risk capital. We will first give a brief outline of the whole model, we refer to as *risk adjusted portfolio management*. The risk of capital investment is being calculated in a multilevel process: We start with a complete portfolio valuation, followed by identifying exposures and aggregating them to exposure classes. Afterwards suitable benchmark indices are mapped on it. Their probability distribution is used to specify market risks of the exposures, which leads to a Value at Risk VaR_P of the portfolio. The absolute value of market risks is supplemented by default risks and risks of misevaluation for special assets, e.g.

- a default risk for corporate or structured bonds R_{D_1}, a default risk for mortgage loans R_{D_2},
- an evaluation risk for strategic participations R_{E_1}, an evaluation risk for property (real estate) R_{E_2},
- a concentration risk R_C for huge proportions of single securities (measuring additional volatility relative to its benchmark) or for accumulations of issuers (risk of issuer default).

Altogether $R_{AP} = |VaR_P| + R_C + R_{D_1} + R_{D_2} + R_{E_1} + R_{E_2}$ is the total risk from the asset portfolio. To get a complete risk balance, we aggregate other risks outside the asset management, e.g. for an insurance company the

[9] As an extension Jorion (2006) presents the delta-gamma approximations, that involve additionally convexity.

underwriting risks R_{UW} and operational risks R_{OP}. These general risk fields have to be aggregated in a certain way, e.g. by copulas on their marginal distributions, or, as proposed by the CEIPOS (2008), by a correlation approach on their absolute measures:[10]

$$R_{tot} := \sqrt{[R_{AP}, R_{OP}, R_{UW}] \begin{bmatrix} 1 & \rho_{AP,OP} & \rho_{AP,UW} \\ \rho_{OP,AP} & 1 & \rho_{OP,UW} \\ \rho_{UW,AP} & \rho_{UW,OP} & 1 \end{bmatrix} \cdot \begin{bmatrix} R_{AP} \\ R_{OP} \\ R_{UW} \end{bmatrix}}, \tag{4}$$

where $\rho_{i,j}$ are pairwise correlations between risk categories i, j. We call R_{tot} the total *risk result*.

The whole approach could be used as a multidimensional controlling tool: With focus on the balance sheet (the profit and loss account, respectively) one can measure a maximum loss of that portfolio of assets, that may lead to amortization, to a release of (hidden) reserves or a deficit in the P&L. This will be called *balance risk*. On the other hand, with a long-term economic view on a market value balance one can measure a maximum loss to the surplus of the liabilities (net equity). Here, the market variation risk of an asset portfolio is regarded relative to contractual commitments, especially in terms of an interest rate risk and credit risk caused by different durations of assets and liabilities. This will be called *ALM risk*. Last but not least one can measure the risk of falling short of regulatory requirements, that is not to cover the actuarial reserve fund by listed securities or to fail the solvency capital requirements in a 'solvency II world' (*regulatory risk*). Interpreting in each of these dimensions the risk result as capital requirement, one has to assure possessing a total *risk capital* C_{tot} at least of that amount. Else, a risk reduction has to be put on immediately, to reach the area of safety margins. The risk capital of an insurance company can be composed of

- coverage capital C_{AP} from the asset portfolio (unrealized gains, (hidden) reserves),
- coverage capital C_{CE} from the contractual commitments (actuarial reserves, discount reserves) and equity capital in a wider sense (including revenue reserves),
- coverage capital C_{PL} from a forecast of the P&L (current/ extraordinary income above target earnings),

whereby $C_{tot} = C_{AP} + C_{CE} + C_{PL}$. Management rules for the asset allocation can be deduced from the proportion of the total risk result R_{tot} and risk capital C_{tot},

$$Q^{\cdot} = \frac{R^{\cdot}_{tot}}{C^{\cdot}_{tot}}, \quad (\cdot \text{ labels the risk dimension}) \tag{5}$$

which should be significantly lower than 100%, so that even in the event of shortfall the budget is most likely not totally consumed.[11]

In the following we come back to the VaR approach for the market risk of assets and focus on an easy implementation of the delta-normal-method. Suppose a broad portfolio of m investments, consisting for the most part of several stocks, multifaceted fixed income securities or investment funds[12] with different denominations. Options, futures and other financial derivatives are kept only in a negligible amount. After marking-to-market all assets $a_1, \ldots, a_m$, the exposures of each asset according to the following general exposure classes have to be identified:

- equity exposures: equity subclasses $(e_1, \ldots, e_{d_1})$ for different economic areas, the corresponding exposures $(w_1, \ldots, w_{d_1})$ are measured as the effective (market valued) investments.
- interest rate exposures: subclasses $(e_{d_1+1}, \ldots, e_{d_1+d_2})$ for different currency areas, the corresponding exposures $(w_{d_1+1}, \ldots, w_{d_1+d_2})$ are measured as the basis point value (bpv) of the assigned securities.[13]

[10] That presumes implicitly, that all risk results are equally scaled standard deviations of normally distributed random variables.

[11] Based on a VaR approach, we assess only a maximum loss to a certain probability, but neither an expectation of such a shortfall nor the maximum possible loss. That is why (depending on the confidence level) not the full amount of risk capital should be on risk. On the other hand, a certain degree of risk is eligible, to achieve a higher expected yield from available investment capital.

[12] Investment funds have to be considered as their single securities. Structured products should be decomposed in their building blocks, if possible.

[13] $w_j = \frac{d_j^{mod} \cdot a_j}{100}$ (in bp), for a single bond with current value a_j and d_j^{mod} as its modified duration ($j = d_1 + 1, \ldots, d_1 + d_2$). For long-positions w_j has a negative sign, since an increase of the corresponding yield curve decreases the fair value.

- credit spread exposures: subclasses $(e_{d_1+d_2+1}, \ldots, e_{d_1+d_2+d_3})$ for different rating classes respectively types of coverage, the corresponding exposures $(w_{d_1+d_2+1}, \ldots, w_{d_1+d_2+d_3})$ are measured as the bpv.[14]
- currency exposures: subclasses $(e_{d_1+d_2+d_3+1}, \ldots, e_d)$, $d = \sum_{i=1}^{4} d_i$, for different denominations, the corresponding exposures $(w_{d_1+d_2+d_3+1}, \ldots, w_d)$ are measured as the effective investments in home currency.

Table 1 represents a real-world implementation, which will be referred to in the remainder of this article. We adopt the perspective of an European investor (EUR as home currency). A single security can contribute exposures to several classes.[15] To obtain a risk measure for each exposure w_i or each class e_i, an applicable benchmark b_i is matched to it. The benchmarks, also called *risk factors* in the pertinent literature, are selected liquid market indices.[16] Via the moments of the benchmark return series $\{X_{i,t}\}_{t=1,\ldots,n}$ the variation of the exposure $w_{i,t}$ shall be estimated later. The general benchmark conception is:

- equity benchmarks $(b_1, \ldots, b_{d_1})$: MSCI equity indices (total return) in local currency
- interest rate benchmarks $(b_{d_1+1}, \ldots, b_{d_1+d_2})$: Swap rates (5 years),[17] in bp
- credit spread benchmarks $(b_{d_1+d_2+1}, \ldots, b_{d_1+d_2+d_3})$: JP Morgan asset swap spread in bp
- currency benchmarks $(b_{d_1+d_2+d_3+1} \ldots, b_d)$: ECB exchange rates, expressed as $\frac{EUR}{x}$ (home currency per foreign unit, price quotation)

A detailed view on all subsequently analysed exposure classes and matched benchmarks[18] gives table 1. The exposure classes require specific return types of benchmark prices $\{P_{i,t}\}_t$: *Arithmetic returns* $X_{i,t} = (P_{i,t} - P_{i,t-1})/P_{i,t-1}$ (alternatively *log-returns* $X_{i,t} = \ln P_{i,t} - \ln P_{i,t-1}$) have to be applied to equities ($i = 1, \ldots, 5$) and currencies ($i = 17, \ldots, 30$), since we measure performances of market values. Differences of prices between time units, so called *diff-returns* $X_{i,t} = P_{i,t} - P_{i,t-1}$, have to be applied to interest rates and credit spreads ($i = 6, \ldots, 16$), since we measure how many times the bpv damps the market value.

In the next step from each benchmark return series $X_{i,1}, \ldots, X_{i,n}$ the (empirical) expected returns μ_i and standard deviations σ_i as well as pairwise correlations $\rho_{i,j}$ ($i \neq j$) are estimated. The basic risk model assumes the returns $X_{i,t}$ to be serially independent and in each period of time identically distributed (iid), following a normally distributed random variable (rv) with expectation μ_i and variance σ_i^2 (henceforth: $\mathcal{N}\left(\mu_i, \sigma_i^2\right)$). A slight modification fixes $\mu_i = 0$ for all i, leading to $\mathcal{N}\left(0, \sigma_i^2\right)$ distributed benchmark yields.[19]

We start univariately measuring the maximum loss, that is not exceeded with a given probability $1 - \alpha$ in a target horizon τ (e.g. $\alpha = 1\%$ by default, $\tau = 10$ days or 3 months), which is the $VaR_i^{1-\alpha,\tau}$ for the benchmark return X_i and exposure w_i (for a fixed point t_0; we omit time subscript whenever possible). In one period it equals the α-quantile $\tilde{u}_\alpha$ of a $\mathcal{N}\left(0, \sigma_i^2\right)$ rv, which can be expressed as the product of the volatility σ_i and the standard normal quantile u_α due to features of the normal distribution. Extending this result over τ units of time, one benefits from the assumption of iid periodical returns $X_{i,t_0+1}, \ldots, X_{i,t_0+\tau}$ and the τ-period volatility is σ_i multiplied by square root of τ. Altogether the α-quantile of the τ-period return has the form

$$\tilde{u}_\alpha^{(\tau)} = u_\alpha \sqrt{\tau}\, \sigma_i =: \nu_i^{(1-\alpha,\tau)}, \tag{6}$$

also called *risk factor* for the exposure amount w_i. Thus, the absolute of the product $\nu_i^{(1-\alpha,\tau)} w_i =: VaR_i^{(1-\alpha,\tau)}$ is the maximum (undiversified) loss or Value at Risk of the exposure w_i in the period $(t_0, t_0 + \tau]$ with probability $1 - \alpha$.

[14] The credit spread to swap curve is modelled here. Again, a long-exposure w_j shows a negative sign, since a spread expansion decreases the bond value.

[15] For example an US-American corporate bond causes an interest rate exposure, a credit spread exposure as well as currency exposure.

[16] Our general data source is Bloomberg. A special thanks is addressed to the index data providers MSCI Barra, JP Morgan and Merrill Lynch as well as to Bloomberg.

[17] The 5-year segment corresponds to the average fixed income duration for many insurers.

[18] Some benchmarks are self created, signed as 'synthetic', according to allocation purposes:
- CredSta: global spread index for (quasi) government guaranteed, 75% Euro +20% US +5% Japan Govt. to Swap Spread (5yr.)
- CredSwa: spread index for pfandbriefe/ covered bonds, 5 year PEX yield to 5 year Swap rate
- CurrEUR: index constant 1 (aim: showing EUR exposure, but having no risk in home currency)
- CurrEM: performance index for all emerging markets currencies ('EM-Dollar' (EMD), exchange rate as EUR/EMD), modelled as difference of periodic returns from MSCI EM index (in EUR) and MSCI EM index (local).

[19] For daily data this is an adequate simplification, for longer periods like monthly returns this corresponds to a conservative left shift of the cdf (in general $\mu_i > 0$, else prefer riskless investment).

ID	exposure class synonym	amount	ID	benchmark (BM) name/ description
e_1	EquEUR	w_1	b_1	MSCI Daily TR Gross EMU Local
e_2	EquexEUR	w_2	b_2	MSCI Daily TR Gross Europe ex EMU Local
e_3	EquNA	w_3	b_3	MSCI Daily TR Gross North America Local
e_4	EquAsP	w_4	b_4	MSCI Daily TR Gross (Asia) Pacific Local
e_5	EquEM	w_5	b_5	MSCI Emerging Markets (Free) Local
e_6	RateEUR	w_6	b_6	EUR Swap annual (30/360), 5 year
e_7	RateUSD	w_7	b_7	USD Swap annual (30/360), 5 year
e_8	RateEUR	w_8	b_8	JPY Swap annual (act/365), 5 year
e_9	CredSta	w_9	b_9	synthetic BM, models Spread Govt. to Swap
e_{10}	CredSwa	w_{10}	b_{10}	synthetic BM, models Spread all Swap-equivalent
e_{11}	CredAAA	w_{11}	b_{11}	JP Morgan Credit Index AAA Asset Swap Spread
e_{12}	CredAA	w_{12}	b_{12}	JP Morgan Credit Index AA Asset Swap Spread
e_{13}	CredA	w_{13}	b_{13}	JP Morgan Credit Index A Asset Swap Spread
e_{14}	CredBBB	w_{14}	b_{14}	JP Morgan Credit Index BBB Asset Swap Spread
e_{15}	CredEM	w_{15}	b_{15}	JP Morgan EMBI Global Divers. Sov. Spread
e_{16}	CredHY	w_{16}	b_{16}	Merrill Lynch HY US BB-B (Spread to US-Swap)
e_{17}	CurrEUR	w_{17}	b_{17}	synthetic BM, set constant 1 (home currency)
e_{18}	CurrGBP	w_{18}	b_{18}	ECB Euro Exchange Ref. Rate as EUR/GBP
e_{19}	CurrCHF	w_{19}	b_{19}	ECB Euro Exchange Ref. Rate as EUR/CHF
e_{20}	CurrSEK	w_{20}	b_{20}	ECB Euro Exchange Ref. Rate as EUR/SEK
e_{21}	CurrDKK	w_{21}	b_{21}	ECB Euro Exchange Ref. Rate as EUR/DKK
e_{22}	CurrNOK	w_{22}	b_{22}	ECB Euro Exchange Ref. Rate as EUR/NOK
e_{23}	CurrUSD	w_{23}	b_{23}	ECB Euro Exchange Ref. Rate as EUR/USD
e_{24}	CurrCAD	w_{24}	b_{24}	ECB Euro Exchange Ref. Rate as EUR/CAD
e_{25}	CurrJPY	w_{25}	b_{25}	ECB Euro Exchange Ref. Rate as EUR/JPY
e_{26}	CurrAUD	w_{26}	b_{26}	ECB Euro Exchange Ref. Rate as EUR/AUD
e_{27}	CurrNZD	w_{27}	b_{27}	ECB Euro Exchange Ref. Rate as EUR/NZD
e_{28}	CurrSGD	w_{28}	b_{28}	Bloomberg exchange rate as EUR/SGD
e_{29}	CurrHKD	w_{29}	b_{29}	Bloomberg exchange rate as EUR/HKD
e_{30}	CurrEM	w_{30}	b_{30}	synthetic BM for exchange rate EUR/EMD

Table 1: Exposure classes e_i of amount w_i and corresponding benchmarks b_i.

Adopting this to the whole exposure vector $\mathbf{w} = (w_1, \ldots, w_d)$ and vector of risk factors $\nu = \left(\nu_1^{1-\alpha,\tau}, \ldots, \nu_d^{1-\alpha,\tau}\right)$ a simple factor-based model with the risk vector $\mathbf{VaR} = \left(VaR_1^{(1-\alpha,\tau)}, \ldots, VaR_d^{(1-\alpha,\tau)}\right)$ results,

$$\mathbf{VaR} = \nu \odot \mathbf{w}\text{ , with } \odot \text{ as elementwise vector multiplication.} \tag{7}$$

A scalar product (of absolutes) would quantify the undiversified risk of the portfolio, $VaR_P^{(undiv)} = |\nu \mathbf{w}'| = \sum_{i=1}^{n} \left|VaR_i^{(1-\alpha,\tau)}\right|$, that neglects diversification benefits from the exposures. Of course, an adequate risk measure has to include the dependence of its risk factors or exposures.

The total portfolio return equals $w_1 X_1 + w_2 X_2 + \ldots + w_d X_d = \mathbf{w}\mathbf{X}' \sim \mathcal{N}\left(0, \sigma_P^2\right)$ for invested exposures $\mathbf{w}$ with normal yields $\mathbf{X}$ according to their benchmarks. With the $d \times d$ correlation matrix $\mathbf{\Omega} = \left(\rho_{i,j}\right)_{i,j}$ it immediately follows for the (monetary) variance σ_P^2 of the portfolio P:

$$\sigma_P^2 = \sum_{i=1}^{d} \sum_{j=1}^{d} w_i w_j \rho_{i,j} \sigma_i \sigma_j = (\sigma \odot \mathbf{w})\, \mathbf{\Omega}\, (\sigma \odot \mathbf{w})' \tag{8}$$

Using property (6) we get the α-quantile of the τ-period portfolio return, in other words the $VaR_P^{(1-\alpha,\tau)}$ for the portfolios market price risk:

$$VaR_P^{(1-\alpha,\tau)} = u_\alpha \sqrt{\tau}\, \sigma_P \tag{9}$$

Alternatively, the risk measure $VaR_P^{(1-\alpha,\tau)}$ can be expressed as:

$$\begin{aligned} VaR_P^{(1-\alpha,\tau)} &= u_\alpha\sqrt{\tau}\sqrt{(\sigma \odot \mathbf{w})\,\mathbf{\Omega}\,(\sigma \odot \mathbf{w})'} \\ &= -\sqrt{\left(\underbrace{(u_\alpha\sqrt{\tau}\sigma_i)}_{=\nu_i} w_i\right)_{i=1,\dots,d} \mathbf{\Omega} \left(\underbrace{(u_\alpha\sqrt{\tau}\sigma_i) w_i}_{=VaR_i}\right)'_{i=1,\dots,d}} \\ &= -\sqrt{\mathbf{VaR}\,\mathbf{\Omega}\,\mathbf{VaR}'} \end{aligned} \tag{10}$$

Summing up, the basic risk model is computationally easy to implement by a matrix multiplication of stressed exposures and correlations. It requires current market values and their first derivations as exposures, benchmarked by appropriate market indices that reflect the variation of exposures and produce risk factors. The chosen risk measure is the Value at Risk $VaR_P^{(1-\alpha,\tau)}$ of the portfolio,[20] estimating a maximum loss over a target horizon τ to a certain confidence $1-\alpha$. On the other hand, a number of criticisms have to be kept in mind:

1. Empirically the return distributions are often asymmetric with fat tails. This is not captured by the proposed model, the VaR restricts the return distribution just to one normal quantile in the left tail. The extent of outliers is not directly described and inadequately modelled due to normal tails.
2. The method insufficiently measures the risk of nonlinear instruments.
3. Besides the assumption of normality for each exposure class, the adjustment to other periods by the square root of time factor assumes that returns are serially independent and identically distributed, which is queried in the next paragraph.

2.2 Testing model assumptions

In the following we are getting granular on the parametric VaR approach by testing its model assumptions. We apply the basic risk model univariately to each exposure class e_i (with amount w_i) respectively their benchmark b_i, $i = 1, \dots, 30$, as defined in table 1. Remember the premises on their returns $\{X_{i,t}\}_t$:

A. the yields of each benchmark index are normally distributed,

B. the returns follow each period the same distribution (identically distributed),

C. the returns of each index are serially independent (independently distributed).

The first attribute allows the simple transformation of volatilities to quantiles $VaR_i^{(1-\alpha,\tau)}$ in a factor-based setup and the multivariate risk aggregation to $VaR_P^{(1-\alpha,\tau)}$ with a Gaussian portfolio return. The second and the third point are necessary for transforming the one-period VaR to any horizon τ. The last one legitimizes the unconditional volatility estimation. In short, we decompose and test the hypothesis of iid normal returns $X_{i,1}, \dots, X_{i,n}$ for the benchmarks b_i, $i = 1, \dots, 30$ univariately. One example of each exposure type is focussed especially: the equity index *EquNA*, the 5-year swap rate *RateUSD*, the government to swap spread *CredSta* and the exchange rate EUR/USD (*CurrUSD*).

A. Testing normal returns

We start testing the normal hypothesis with monthly returns of the benchmark indices (monthly closing prices) from January 29, 1999 to December 29, 2006 (95 return points), using arithmetic returns for equity indices and exchange rates and diff-returns for interest rates and credit spreads (standard conception of basic risk model). In the following the null hypothesis

$$H_0: \text{ sample is normally distributed}$$

[20] In the proceeded normal model the risk measure can be straightforward transformed to the *expected shortfall*, that measures the expected loss $ES_P^{(1-\alpha,\tau)}$ in the case of a loss exceeding the $VaR_P^{(1-\alpha,\tau)}$. It can be derived that $ES_P^{(1-\alpha,\tau)} = -\frac{\phi(u_\alpha)}{1-\alpha}\sqrt{\tau}\sigma_P = -\frac{\phi(u_\alpha)}{u_\alpha(1-\alpha)}VaR_P^{(1-\alpha,\tau)}$, i.e the ES can be expressed as a real multitude (> 1) of the VaR depending on the confidence level.

is tested on the 5% level of significance via the χ^2-goodness of fit test, the Kolmogorov-Smirnov (K-S) test, the Anderson-Darling (A-D) test and the Jarque-Bera (J-B) test. The statistical tests are graphically supplemented by quantile-quantile-plots (qq-plots) of ordered sample returns against normal quantiles, provided in figure 1.

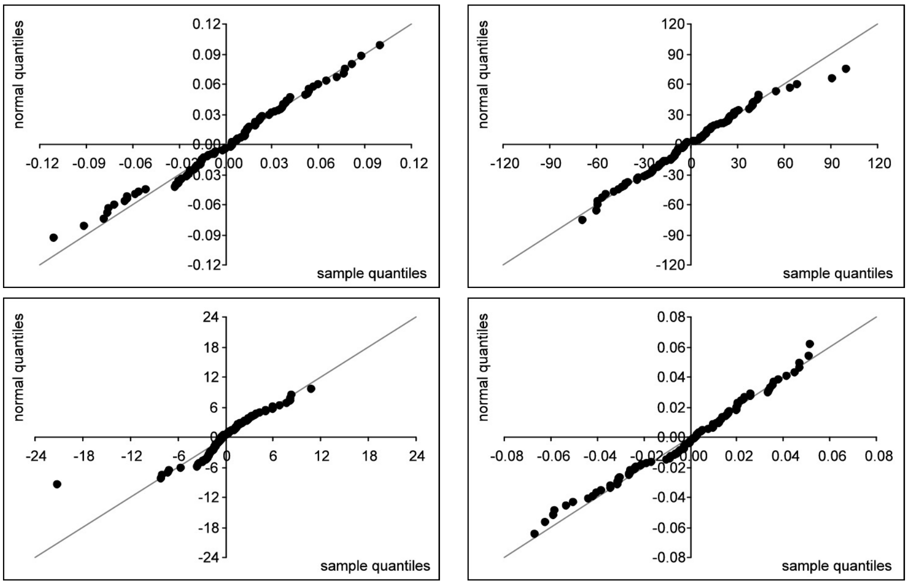


Figure 1: QQ-plots for monthly returns of *EquNA*, *RateUSD*, *CredSta* and *CurrUSD* against quantiles of the normal distribution (from top left to bottom right, line-by-line).

The qq-plots of the monthly benchmark returns indicate a surprising good approximation by a normal distribution, especially for the equity and currency example the points are quite close to the bisecting line. The hypothesis tests come to similar results on the 5% level, presented in figure 2. Most rejections (12 of 29 benchmarks[21]) occur to the J-B test on skewness and excess kurtosis.[22] Within the distance tests, the A-D is the strongest leading to 11 rejections, followed by K-S, where H_0 is 9 times denied. We summarize the testing results per benchmark in the following way:

H_0 is rejected at all, if more than 2 tests are not passed.

This summary test heuristic is chosen quite weak, it leads to 7 overall rejections, that are reported in the tableau, figure 3, aggregated within the general exposure types. From this it follows, that the normal fit is worst for the credit spread indices with 4 rejections, apart from that a satisfactory approximation is observed.

The normal hypothesis tests are repeated with the return types arithmetic, diff-returns and log-returns applied to all benchmark indices successively. The target is getting a qualitative result, what return type works best (for which exposure classes) concerning the normal approximation. The outcomes for the three types are very close to each other,[23] with summary test heuristics similar to figure 3. We conclude in preferring as best return conception:

- log-returns for equities and currencies,
- diff-returns for interest rates and credit spreads.

[21] *CurrEUR* is removed from hypothesis testing, because the riskless class was set constant and only introduced for exposure completeness.

[22] In a basic statistical evaluation on all benchmarks, it stands out that the centered third moments and fourth moments (minus 3) deviate from zero. More than a half of the time series tend to be leptokurtic. Most benchmark returns are skewed to its loss tail.

[23] For lack of space we do not reprint the full statistics. Concerning equity returns there are marginal distinctions in the normal fit: diff-returns deliver worst results, log-returns may be preferred because of their convenient features. For interest rates next to smaller differences (arithmetic returns worst) modelling errors may arise to arithmetic and esp. log-returns when rates become zero or negative; that's why diff-returns are preferred. The same technical failure is quite likely for credit spreads due to zero or negative spreads (in the higher ratings), so that they should be modelled as diff-returns. For currencies we observed only marginal distinctions amongst the conceptions and log-returns are (subjective) favored.

Normal hypothesis tests										
J-B test	**EquEUR**	**EquexEUR**	**EquNA**	**EquAsP**	**EquEM**	**RateEUR**	**RateUSD**	**RateJPY**	**CredSta**	**CredSwa**
-> test statistic	12.473	14.031	1.039	0.019	0.999	2.249	3.499	98.794	240.612	8.182
-> critical value	5.991	5.991	5.991	5.991	5.991	5.991	5.991	5.991	5.991	5.991
Chi^2 test										
-> test statistic	11.842	16.263	14.368	2.790	3.842	14.579	4.895	18.789	19.210	14.790
-> critical value	16.919	16.919	16.919	16.919	16.919	16.919	16.919	16.919	16.919	16.919
K-S test										
-> test statistic	0.771	1.150	0.642	0.487	0.516	0.814	0.577	1.184	1.346	0.595
-> critical value	0.890	0.890	0.890	0.890	0.890	0.890	0.890	0.890	0.890	0.890
A-D test										
-> test statistic	1.061	1.884	0.458	0.259	0.306	0.456	0.293	2.088	2.322	0.484
-> critical value	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752
summary										
-> rejection	no	yes	no	no	no	no	no	yes	yes	no

J-B test	**CredAAA**	**CredAA**	**CredA**	**CredBBB**	**CredEM**	**CredHY**	**CurrEUR**	**CurrGBP**	**CurrCHF**	**CurrSEK**
-> test statistic	2.945	132.924	54.328	57.346	5.209	29.525		10.278	2.280	1.172
-> critical value	5.991	5.991	5.991	5.991	5.991	5.991		5.991	5.991	5.991
Chi^2 test										
-> test statistic	1.105	11.632	15.632	31.000	13.526	12.053		7.000	11.421	4.474
-> critical value	16.919	16.919	16.919	16.919	16.919	16.919		16.919	16.919	16.919
K-S test										
-> test statistic	0.575	1.112	1.265	1.271	0.916	0.887		0.591	0.781	0.547
-> critical value	0.890	0.890	0.890	0.890	0.890	0.890		0.890	0.890	0.890
A-D test										
-> test statistic	0.286	1.712	2.530	2.309	0.938	0.908		0.447	0.634	0.353
-> critical value	0.752	0.752	0.752	0.752	0.752	0.752		0.752	0.752	0.752
summary										
-> rejection	no	yes	yes	yes	no	no		no	no	no

J-B test	**CurrDKK**	**CurrNOK**	**CurrUSD**	**CurrCAD**	**CurrJPY**	**CurrAUD**	**CurrNZD**	**CurrSGD**	**CurrHKD**	**CurrEM**
-> test statistic	0.272	4.829	1.961	2.493	18.245	0.347	9.692	0.022	0.744	5.758
-> critical value	5.991	5.991	5.991	5.991	5.991	5.991	5.991	5.991	5.991	5.991
Chi^2 test										
-> test statistic	4.474	16.684	11.211	11.842	9.526	5.105	6.790	14.369	4.684	8.474
-> critical value	16.919	16.919	16.919	16.919	16.919	16.919	16.919	16.919	16.919	16.919
K-S test										
-> test statistic	0.538	0.972	0.846	0.513	0.938	0.446	0.429	0.467	0.777	0.783
-> critical value	0.890	0.890	0.890	0.890	0.890	0.890	0.890	0.890	0.890	0.890
A-D test										
-> test statistic	0.218	0.926	0.573	0.449	1.593	0.188	0.407	0.172	0.227	0.371
-> critical value	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752
summary										
-> rejection	no	no	no	no	yes	no	no	no	no	no

Figure 2: Goodness of fit tests against normality for benchmark indices from table 1. Rejected tests are grey highlighted. The summary test statistic rejects, if more than 2 single tests are rejected.

number of rejected normal hypotheses in summary (standard conception)	**7**
-> thereof equities	1
-> thereof interest rates	1
-> thereof credit spreads	4
-> thereof currencies	1

Figure 3: Test summary for Gaussian returns in the standard conception.

Leaving the monthly base, we now reinvestigate if the previous results hold for daily returns. Our benchmark data is refined to daily closing prices from January 4, 1999 to December 29, 2006 (2061 returns). This time the qq-plots in figure 4 present a different impression. The normal fit seems to be very bad: on all four examples strong, systematic deviations can be observed. The sample tails are in most instances much heavier than a normal distribution and asymmetric behavior of gains and losses is to be noticed in some extent.[24]

These results are statistically confirmed by testing the hypothesis H_0 with the four test statistics. All singular tests are rejected on a 5% level for all benchmark return series. Mostly the test statistic is far apart from the critical value.[25] The summarizing heuristic, as shown in figure 5 is also identical for all return types, neglecting H_0 for all benchmarks. The normal distribution of empirical daily returns is clearly rejected.

[24] An extension of this analysis to exponential quantiles exhibited for some examples, that return data (with sample gains and losses separately fitted) decays slightly slower than tails of the exponential distribution. Another argument for asymmetry can be obtained by comparing the gains with the losses, e.g. by plotting the empirical quantiles of X_t against $-X_t$.

[25] Interested readers are welcome to ask us for the complete test results, we will gladly provide analogous to figure 2.

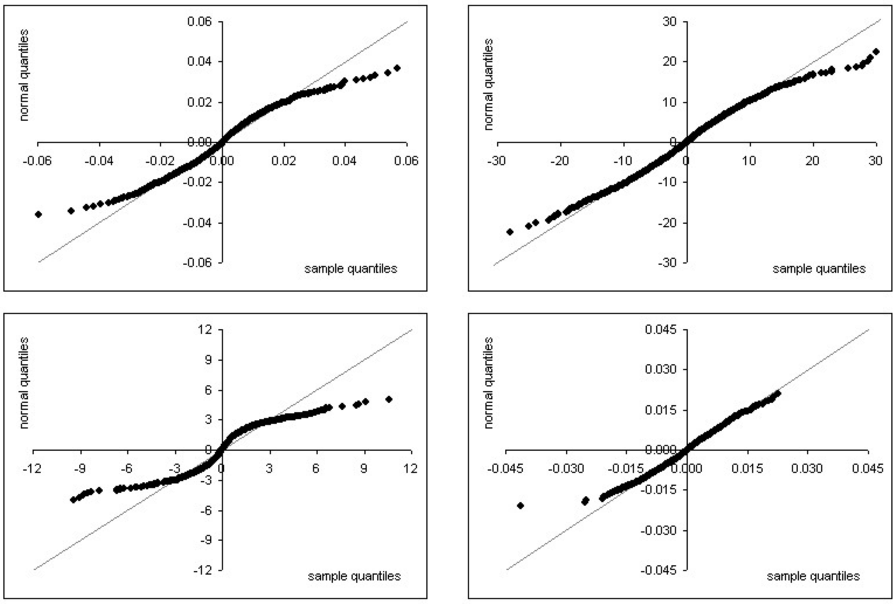


Figure 4: QQ-plots for daily returns of *EquNA*, *RateUSD*, *CredSta* and *CurrUSD* against normal quantiles (from top left to bottom right, line-by-line).

number of rejected normal hypotheses in summary (for each return type)	29
-> thereof equities	5
-> thereof interest rates	3
-> thereof credit spreads	8
-> thereof currencies	13

Figure 5: Test summary heuristic, which is identical for all types of daily returns, rejecting each normal hypothesis.

The testing results on daily and monthly normal returns are contrary, a temporal closedness of the risk model and the assumed iid returns are queried furthermore. But are daily performances indeed so much worse against monthly returns regarding the normal approximation? We try to answer this by enlarging the monthly data base to price series from January 31, 1991 to March 31, 2008 (206 return points).[26] Replaying the monthly analysis, a clearly worse normal fit results compared to the horizon 1999 to 2006. The number of total rejections increases from 7 to 17. Especially for equity benchmarks (4 rejections, former 1) and credit spreads (7 rejections, former 4) normality is rejected mostly. Possible reasons may be a stronger robustness of that sample and tighter confidence bands, that improve statistical quality even if short- and long-term data are identically distributed. The homogenous normal distribution has to be challenged on the long-term monthly sample. Because different sample sizes and periodicities imply different results, the temporal closedness and the normality of the risk model has to be denied after all. Furthermore, concluding on the J-B tests, a heavy tailed, asymmetric return distribution should be imposed instead.

B. Testing identically distributed returns

The risk model is based on the assumption that returns are subject to the same probability distribution each period. This can be proved for each benchmark by dividing the time series in several subsamples and comparing their distributions pairwise. Similar techniques as in the previous paragraph can be applied by assuming that the (unknown)

[26] Some benchmark indices have to be approximated by similar reference indices, most important: *RateEUR*: DEM Swap before 1999; *CredSwa*: 100% US govt. to swap before 1996; *CredAAA,...*: Merrill Lynch US Corp. Spreads before 1999; *Curr...*: exchange rate DEM/x converted via EUR-fixing (EUR introduction on January 1, 1999).

distribution of subsample A is the hypothetical distribution of subsample B. The null hypothesis

$$H_0: \text{ subsample } A \text{ is identically distributed as subsample } B$$

is tested on the 5% level of significance via the two-sample Kolmogorov-Smirnov (K-S) test and the two-sample χ^2-goodness of fit test.[27] Basically, identical distributions need to have identical moments, that is proved with the t-test for expectations $\mu_A = \mu_B$ (restricted to $\sigma_A^2 = \sigma_B^2$) and the F-test for variances $\sigma_A^2 = \sigma_B^2$. The formal tests may be graphically supplemented by qq-plots of subsample returns.

	Identical distribution tests									
K-S test	**EquEUR**	**EquexEUR**	**EquNA**	**EquAsP**	**EquEM**	**RateEUR**	**RateUSD**	**RateJPY**	**CredSta**	**CredSwa**
critical value (c-value)	*1.36*	*1.36*	*1.36*	*1.36*	*1.36*	*1.36*	*1.36*	*1.36*	*1.36*	*1.36*
1999-2000 vs. 2001-2002	1.6330	1.6330	1.6330	1.2247	0.8165	1.4289	1.8371	1.2247	1.4289	1.4289
1999-2000 vs. 2003-2004	1.2247	0.8165	1.6330	1.2247	1.4289	1.2247	0.4082	1.6330	0.6124	1.0206
1999-2000 vs. 2005-2006	1.4289	1.2247	1.6330	1.4289	1.2247	0.8165	0.6124	1.6330	1.0206	1.2247
2001-2002 vs. 2003-2004	1.6330	2.2454	2.0412	2.0412	1.4289	1.0206	1.4289	1.2247	1.4289	1.0206
2001-2002 vs. 2005-2006	2.4495	2.2454	1.8371	2.0412	1.6330	1.8371	2.0412	1.8371	1.8371	0.8165
2003-2004 vs. 2005-2006	1.0206	1.0206	1.0206	0.8165	1.0206	1.2247	0.8165	1.8371	0.8165	0.8165
1999-2002 vs. 2003-2006	2.1651	2.1651	2.3094	2.0207	1.4434	1.0104	1.2990	1.5877	1.5877	0.8660
even vs. odd months	1.7321	2.0207	1.0104	1.1547	1.2990	0.7217	0.8660	1.2990	1.1547	1.7321
-> # rejected	6	5	6	4	4	2	3	5	4	2
Chi^2 test										
c-value subsequent	*7.8147*	*7.8147*	*7.8147*	*7.8147*	*7.8147*	*7.8147*	*7.8147*	*7.8147*	*7.8147*	*7.8147*
1999-2000 vs. 2001-2002	20.3333	15.6667	9.0000	6.6667	6.0000	7.0000	9.0000	13.6667	6.3333	11.0000
1999-2000 vs. 2003-2004	9.3333	3.3333	18.0000	9.0000	12.3333	5.6667	1.0000	9.0000	3.0000	3.0000
1999-2000 vs. 2005-2006	10.3333	3.0000	39.0000	4.3333	8.0000	5.0000	0.3333	6.0000	13.6667	5.6667
2001-2002 vs. 2003-2004	21.0000	21.0000	16.3333	14.3333	6.3333	1.3333	8.3333	6.3333	8.6667	11.0000
2001-2002 vs. 2005-2006	33.0000	23.6667	17.0000	16.6667	8.3333	11.0000	9.3333	16.3333	21.0000	4.3333
2003-2004 vs. 2005-2006	5.0000	3.6667	3.0000	2.3333	5.0000	6.0000	2.3333	18.3333	2.3333	1.6667
c-value subsequent	*14.0671*	*14.0671*	*14.0671*	*14.0671*	*14.0671*	*14.0671*	*14.0671*	*14.0671*	*14.0671*	*14.0671*
1999-2002 vs. 2003-2006	399.6429	69.1714	761.2081	390.9584	400.3714	9.0584	13.9429	22.1429	53.5920	37.5429
even vs. odd months	64.4692	395.5442	31.0857	10.4727	55.3121	40.9000	36.7584	11.3214	19.0857	43.8605
-> # rejected	7	5	7	4	5	2	4	5	5	4
t-test										
c-value subsequent	*2.0129*	*2.0129*	*2.0129*	*2.0129*	*2.0129*	*2.0129*	*2.0129*	*2.0129*	*2.0129*	*2.0129*
1999-2000 vs. 2001-2002	1.9290	1.8691	1.2607	1.7793	0.6117	1.8278	1.6112	0.1883	1.3709	0.2307
1999-2000 vs. 2003-2004	0.1197	0.6639	1.0183	0.4442	0.6468	1.2674	0.0628	0.8055	0.9942	0.2417
1999-2000 vs. 2005-2006	0.6015	1.1116	0.4573	0.7724	0.7020	0.3710	0.2095	1.1473	0.5177	0.4682
2001-2002 vs. 2003-2004	2.1101	2.4938	2.4130	2.7128	1.3162	0.5396	1.3971	1.4337	0.5893	0.0705
2001-2002 vs. 2005-2006	2.7843	3.0616	2.0640	2.9098	1.3408	1.7295	1.9370	1.9615	1.5504	0.2559
2003-2004 vs. 2005-2006	0.4911	0.3740	0.9281	0.4541	0.1016	1.0914	0.0971	0.3867	0.9680	0.1203
c-value subsequent	*1.9855*	*1.9855*	*1.9855*	*1.9855*	*1.9855*	*1.9855*	*1.9855*	*1.9855*	*1.9855*	*1.9855*
1999-2002 vs. 2003-2006	2.0691	2.6453	2.1758	2.3261	1.4448	0.1899	1.2875	1.7352	0.1211	0.3474
even vs. odd months	2.3299	2.2502	0.8649	1.1310	1.0058	0.1819	0.6719	1.0144	1.9082	1.6750
-> # rejected	4	4	3	3	0	0	0	0	0	0
F-test										
max c-value subsequent	*2.3116*	*2.3116*	*2.3116*	*2.3116*	*2.3116*	*2.3116*	*2.3116*	*2.3116*	*2.3116*	*2.3116*
min c-value subsequent	*0.4326*	*0.4326*	*0.4326*	*0.4326*	*0.4326*	*0.4326*	*0.4326*	*0.4326*	*0.4326*	*0.4326*
1999-2000 vs. 2001-2002	0.4307	0.4892	0.6394	1.3626	0.7633	0.8162	0.5277	5.6822	1.1851	1.3247
1999-2000 vs. 2003-2004	1.2607	1.2450	2.7702	2.5385	2.1923	0.8277	0.4594	1.7757	2.0922	0.4663
1999-2000 vs. 2005-2006	3.1639	2.3364	5.0718	1.8229	1.7774	1.8004	1.5853	1.8927	9.0265	1.3109
2001-2002 vs. 2003-2004	2.9268	2.5450	4.3328	1.8631	2.8722	1.0141	0.8707	0.3125	1.7654	0.3520
2001-2002 vs. 2005-2006	7.3454	4.7762	7.9327	1.3379	2.3286	2.2057	3.0042	0.3331	7.6167	0.9896
2003-2004 vs. 2005-2006	2.5097	1.8767	1.8309	0.7181	0.8107	2.1751	3.4504	1.0659	4.3144	2.8114
max c-value subsequent	*1.7838*	*1.7838*	*1.7838*	*1.7838*	*1.7838*	*1.7838*	*1.7838*	*1.7838*	*1.7838*	*1.7838*
min c-value subsequent	*0.5606*	*0.5606*	*0.5606*	*0.5606*	*0.5606*	*0.5606*	*0.5606*	*0.5606*	*0.5606*	*0.5606*
1999-2002 vs. 2003-2006	3.2105	2.6467	4.6652	1.9568	2.2854	1.3202	1.0878	1.0746	3.2014	0.6041
even vs. odd months	0.9659	1.5398	1.1093	1.3883	1.5649	0.7988	1.2828	0.3853	1.9847	1.1230
-> # rejected	6	4	5	2	3	0	2	4	5	2
summary										
-> rejection	yes	yes	yes	yes	yes	no	no	yes	yes	no

K-S test	**CredAAA**	**CredAA**	**CredA**	**CredBBB**	**CredEM**	**CredHY**	**CurrEUR**	**CurrGBP**	**CurrCHF**	**CurrSEK**
-> # rejected	3	4	6	7	7	4		5	5	0
Chi^2 test										
-> # rejected	1	5	6	7	7	5		6	4	5
t-test										
-> # rejected	0	0	0	1	1	1		0	2	0
F-test										
-> # rejected	0	4	6	4	5	5		5	2	3
summary										
-> rejection	no	yes	yes	yes	yes	yes		yes	yes	no

K-S test	**CurrDKK**	**CurrNOK**	**CurrUSD**	**CurrCAD**	**CurrJPY**	**CurrAUD**	**CurrNZD**	**CurrSGD**	**CurrHKD**	**CurrEM**
-> # rejected	1	1	4	3	4	0	0	4	4	4
Chi^2 test										
-> # rejected	3	2	4	4	7	2	2	7	6	6
t-test										
-> # rejected	0	1	1	0	0	0	0	1	1	0
F-test										
-> # rejected	0	3	0	0	3	1	0	2	0	3
summary										
-> rejection	no	no	yes	no	yes	no	no	yes	yes	yes

Figure 6: Tests for identically distributed monthly returns of benchmark indices from table 1.

[27] The test statistics are approximative for large samples. Our K-S implementation measures the maximum difference of steps of the empirical distribution functions of subsamples A and B. The χ^2-test sums up k class-wise differences of relative frequencies. Our class fragmentation is fixed on quantiles of the 'null subsample'. This approach is not symmetrically, i.e. changing the 'null part' may lead to different results. We test both ways and compare the maximum test statistic to the critical value $\chi^2_{k-1;95\%}$.

Again, we work first with the monthly and afterwards with the daily data base from 1999 to 2006 of benchmark returns[28] and analyse the impact of different return types. The subsamples are chosen as:

(i) subsequences of 2 years of data (length: 24 months or 515 days)

(ii) subsequences of 4 years of data (length: 48 months or 1030 days)

(iii) subsequences of even vs. odd months (respectively days alternating)

For point (i) 6 combinations are possible, altogether 8 (pairwise) input series can be arranged and are proved by each of the id-test procedures. On a singular approach (e.g. K-S) the results are summed up in the following way:

H_0 is (preliminary) rejected if more than 3 pairs fail the test.

Former we define a summary rule on all testing approaches:

H_0 is rejected overall if K-S and χ^2-test fail <u>or</u> if more than 2 tests fail.

Starting with monthly returns, figure 6 presents for the first ten benchmark series the detailed test statistics. For the remaining samples only the number of rejected pairs regarding the approaches and the summary heuristic is shown.[29] The χ^2-test statistic seems to be the most severe with 23 rejected benchmark series, followed by the K-S test with 19 rejections. The serial identity of all equity samples is rejected by both tests and the majority of credit spreads fail. Only some currencies as well as the EUR and USD swap rate pass the identity assumption. These results dominate the summary test heuristic, which denies H_0 likewise 19 times. The t-test for identical expectations is only for European equity indices denied. The test of homoscedastic subsamples (F-test) leads to 11 classwise rejections, above all at equity and credit series.

Paying attention which return types fulfill the assumption of identity best, there are slight differences observed. But the stronger influence is again exerted by possible modelling errors at zero/ negative rates or spreads when using arithmetic- and especially log-returns. This leads to the same predominances as derived in the foregoing paragraph and a preferred conception consisting of log-returns for equities and currencies and diff-returns for interest rates and credit spreads.

Switching to daily returns, we start this time with qq-plots of bisected samples for the return series *EquNA*, *RateUSD*, *CredSta* and *CurrUSD*. Figure 7 shows varying results on the exposure classes: The fit to the bisecting line is especially bad for the equity and the credit spread example, where the extent of returns is stronger in the first four years and the distribution seems to differ between the series' subsamples. Some tail deviations are observed for the *CurrUSD* example. A visually quite good quantile-quantile fit appears for the daily US swap rates.

The graphical impression is not fully affirmed by the hypothesis tests: Because of the longer data base and closer non-rejection areas, most pairs of subsamples are rejected by the goodness of fit tests and the F-test, even the USD interest rate. The only exception is the spread series *CredSwa* that scrapes through the K-S test and therewith passes at all. For all benchmarks identical daily expectations (close to zero) are confirmed with the t-test. Figure 8 summarizes the test rejections for the exposure types according to the predefined rules.[30] H_0 is rejected for 28 benchmarks on the daily data base and the periodical identity of the return distribution has to be strongly rejected. This result holds for the different return types and the before preferred return conception can be recommended again.

Concluding on all identical distribution tests, this assumption does not hold for the majority of monthly samples, with a particularly bad fit of equity and credit spread subsamples. On the daily refinement the univariate identity of return series can be denied at all. It immediately follows that a stationary assumption on that return series is inappropriate. Furthermore, since the F-test for identical variances fails in the majority of disjoint subsamples for all benchmarks, a homoscedastic implementation is denied. As a consequence a non-stationary, heteroscedastic financial time series model should be chosen.

[28] For practical reasons (sample size divisible by four) the monthly data base is extended by one return point for January 1999. On the other hand, in the daily samples the first return point (January 5, 1999) is removed.

[29] For lack of space we do not display qq-plots of monthly subsamples. Moreover the small data bases would complicate an interpretation, whether quantile deviations are random or systematic. A good example was the bisected *EquNA* series, where the distance to the bisecting line is especially obvious, indicating a stronger extent of North American equity returns in the first half (1999-2002) of the sample.

[30] Interested readers are welcome to ask us for the complete test results, we will gladly provide analogous to figure 6.

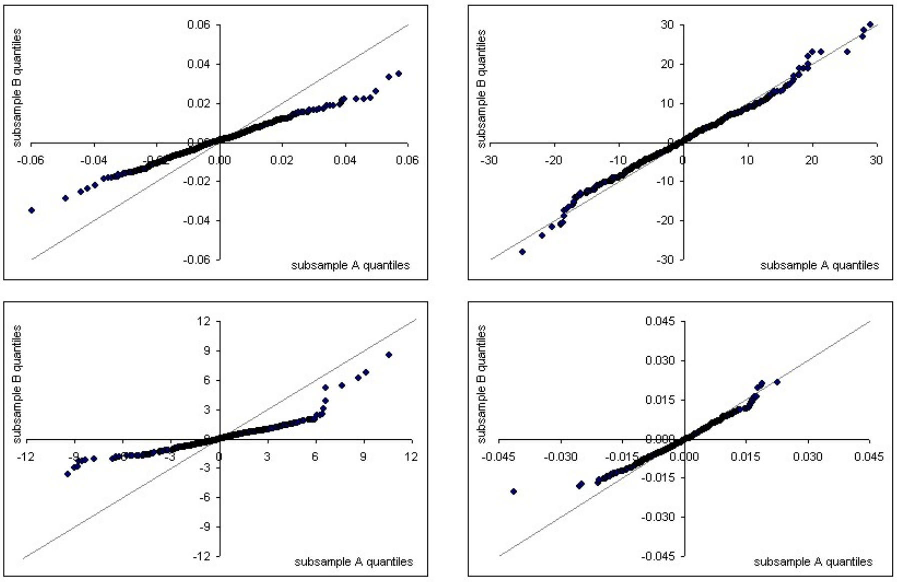


Figure 7: Two-sample qq-plots for daily returns 1999-2002 vs. 2003-2006 of *EquNA*, *RateUSD*, *CredSta* and *CurrUSD* (from top left to bottom right, line-by-line).

number of rejected identity hypotheses in summary (for each return type)	28
-> thereof equities	5
-> thereof interest rates	3
-> thereof credit spreads	7
-> thereof currencies	13

Figure 8: Test summary heuristic concerning the serial identity of daily benchmark returns.

C. Testing serially independent returns

It is a well known fact in mathematical statistics that the consistency of the empirical mean or variance requires a stochastically independent series of random variables with some equal, finite moments. So, serially independent benchmark returns $X_{i,1}, \ldots, X_{i,n}$ $(i = 1, \ldots, d)$ are necessary for volatility estimation as introduced in the basic risk model. A necessary condition for independent realisations is the uncorrelatedness within a time series. This can be checked via correlograms, that depict sample autocorrelations $\rho(h)$ against time lags $h \in \mathbb{N}_0$.[31] To give evidence whether a SACF differs significantly from zero, we develop a criterion with some basic statistical considerations:[32] For an iid series $X_1, \ldots, X_n$

$$\left[\pm \frac{1}{\sqrt{n}}\, t_{n-1;1-\frac{\alpha}{2}}\right] \tag{11}$$

is an approximative $(1-\alpha)$ confidence interval of $\gamma(h)$, $h \in \mathbb{N}_0$.

While a correlogram graphically checks a necessary linear condition, an easy nonlinear extension is to analyse parallel the absolute values of returns (or their power transformations) with their correlogram. Here we can profit from additional criteria: An uncorrelated stationary return series $X_1, \ldots, X_n$ that has autocorrelated absolutes $\{|X_t|\}_t$

[31] Compare section 1. This analysis implicitly requires a stationary stochastic process.

[32] For iid normal distributed rvs $X_1, \ldots, X_n$ with $EX_1 = \mu$ and $Var X_1 = \sigma^2$ unknown, $\left[\bar{X}_n \pm \frac{\hat{\sigma} t_{n-1;1-\frac{\alpha}{2}}}{\sqrt{n}}\right]$ is an exact $(1-\alpha)$ confidence interval for μ, where $t_{m;\alpha}$ denotes the α quantile of a student distribution with m df. Without the normality assumption the confidence interval for μ is approximative. From Chan (2002) we apply, that for an iid series $X_1, \ldots, X_n$ it holds for each $h \in \mathbb{N}_0$ that $\rho(h) \xrightarrow{D} \mathcal{N}\left(0, \frac{1}{n}\right)$, with D denoting the convergence in distribution.

(or power transformations $\{X_t^p\}_t$, $p \in \mathbb{R}$) is serially dependent.[33] Moreover, a nice corollary whether the return distribution is possibly Gaussian follows:[34] If a stationary return sample $X_1, \ldots, X_n$ is serially uncorrelated but their absolutes (or power transformations) are significantly autocorrelated, then a normal distribution cannot hold.

Having the necessary theory, we analyse graphically the serial dependence structure for the reference examples on the 8-year monthly benchmark series,[35] plotting their correlograms in figure 9. The displayed SACFs $\rho_{X_t}(h)$ on returns and their absolutes $\rho_{|X_t|}(h)$ are for the very most lags within the 95% acceptance region. In the *EquNA* example four bars (regarding absolute returns) exceed the upper limit marginally, but no systemtatic structure is obvious. The correlograms give a clear indication that there is no serial dependence in the monthly benchmark returns.[36]

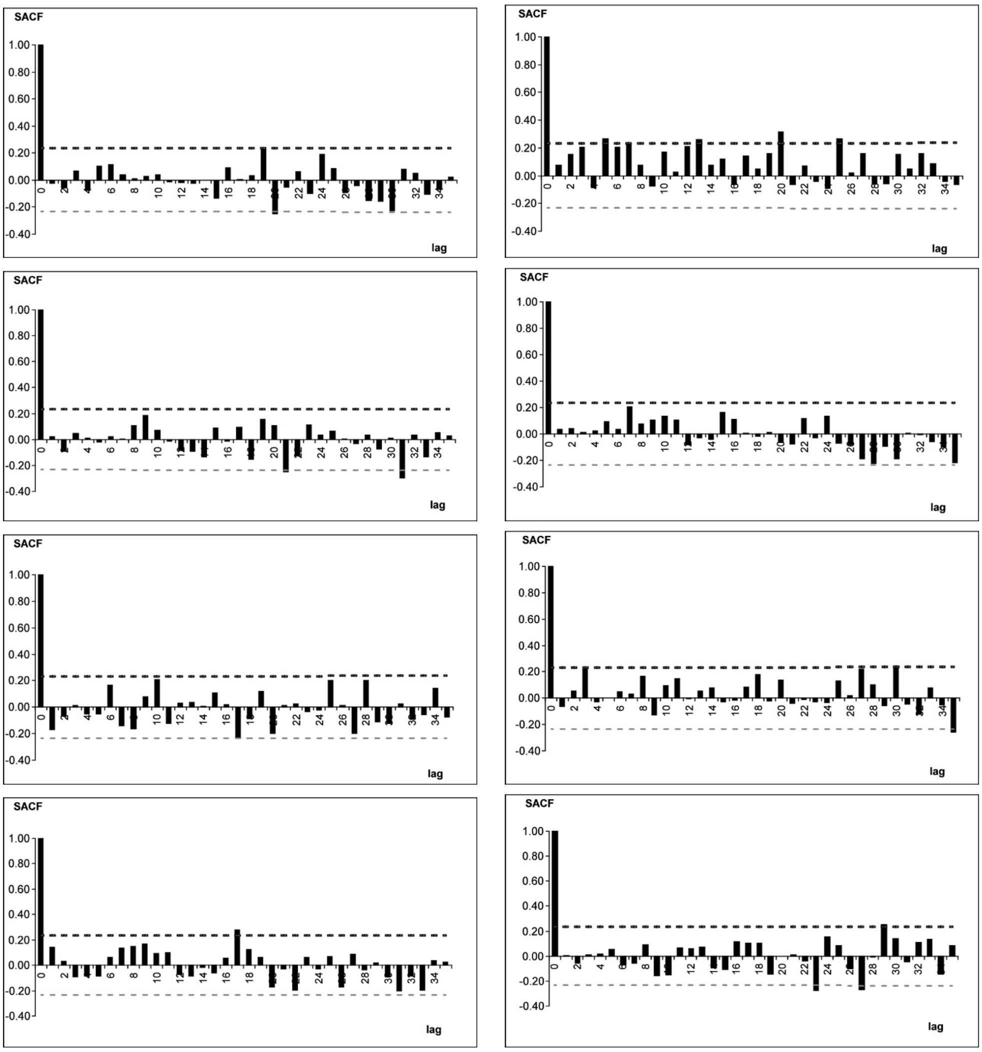


Figure 9: Correlograms for monthly returns (left) and their absolutes (right) of *EquNA*, *RateUSD*, *CredSta* and *Curr-USD* (from top to bottom). The dashed lines define the approximative 95% acceptance region for iid returns.

[33] This follows from the equivalence of independence between rvs X, Y and independence between rvs $f(X), f(Y)$ for continuous functions f, with contraposition on uncorrelatedness of $f(X), f(Y)$.

[34] Since independence and uncorrelatedness are equivalent for normal distributed rvs, this must hold for their continuous functions; compare footnote 33.

[35] We stay in the standard return conception. The other return types are supposed to show the same results regarding independence, because of the aforementioned transformation result.

[36] Several credit spread examples, as *CredSwa*, *CredAAA* or *CredHY*, show somewhat stronger autocorrelations, especially in the first lag.

Validating this result in a formal way, the null hypothesis

$$H_0: \text{ sample is serially independent}$$

is tested on the 5% significance level via a Portmanteau test in the version of Ljung-Box (LB)[37] and via the test statistic of Brock, Dechert, Scheinkman (BDS independence test)[38]. Figure 10 gives an overview of test results concerning all monthly benchmark series. LB accepts for 20 of the 29 indices the serial independence, with no rejections on stock indices and interest rates. Some credit spreads and currencies (e.g. *CurrUSD*) show negative results. The BDS independence test rejects independence for the Euro area and North American equities as well as for the EUR swap and all credit spreads except *CredSta*, but currencies show a good performance. Altogether BDS arrives at 11 rejections. Again a summary test heuristic for both tests is defined:

$$H_0 \text{ is overall rejected if both tests fail.}$$

Following that rule only three rejections are left. But remember, that the small sample size implies rather broad acceptance regions and that the assumption of identically distributed returns is needed.

Independence tests

Portmanteau test	EquEUR	EquexEUR	EquNA	EquAsP	EquEM	RateEUR	RateUSD	RateJPY	CredSta	CredSwa
test statistic LB	19.3814	16.6068	22.3664	19.4862	20.6738	22.3626	19.5769	13.9080	34.1111	70.2480
critical value	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*
-> rejection	no	no	no	no	no	no	no	no	yes	yes
BDS independ. test										
-> rejection	yes	no	yes	no	no	yes	no	no	no	yes
summary										
-> rejection	no	no	no	no	no	no	no	no	no	yes

Portmanteau test	CredAAA	CredAA	CredA	CredBBB	CredEM	CredHY	CurrEUR	CurrGBP	CurrCHF	CurrSEK
test statistic LB	32.0077	21.1892	13.2289	12.6822	23.4166	18.5277		34.2735	17.3151	22.6140
critical value	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*		*31.4104*	*31.4104*	*31.4104*
-> rejection	yes	no	no	no	no	no		yes	no	no
BDS independ. test										
-> rejection	yes	yes	yes	yes	yes	yes		yes	no	no
summary										
-> rejection	yes	no	no	no	no	no		yes	no	no

Portmanteau test	CurrDKK	CurrNOK	CurrUSD	CurrCAD	CurrJPY	CurrAUD	CurrNZD	CurrSGD	CurrHKD	CurrEM
test statistic LB	24.8032	39.3732	31.9135	32.3232	24.5981	21.6853	13.4957	42.6774	37.7751	27.9849
critical value	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*	*31.4104*
-> rejection	no	yes	yes	yes	no	no	no	yes	yes	no
BDS independ. test										
-> rejection	no	no	no	no	no	no	no	no	no	no
summary										
-> rejection	no	no	no	no	no	no	no	no	no	no

Figure 10: Tests for stochastically independent monthly returns of benchmark indices, table 1.

Switching to daily returns on the same horizon, figure 11 presents the correlograms for the benchmark returns $\{X_t\}_t$ and their absolutes $\{|X_t|\}_t$. While again the SACF of original returns (except *CredSta*)[39] shows only a few outliers from the confidence band $[-0.05, 0.05]$, a systematic and significant autocorrelation appears in the absolute returns. The SACF is strictly positive for all lags and declines slowly with the number of lags. A notably high level of absolute autocorrelation arises from the *CredSta* ($\rho_{|X_t|}(h) \geq 0.2$ for $h = 1, \ldots, 50$) and *EquNA* series. As motivated in the introduction, this could be interpreted as a (stationary) long range dependence.[40] With the criteria above, independence of daily returns is hardly requested and the features contradict to a normal distribution.

Once again, the result is statistically supplemented by testing H_0 with the LB and BDS test for all benchmarks. As displayed in figure 12 the assumption of serially independent returns is rejected on the majority of indices. The BDS test denies H_0 for all but one (*CurrHKD*) benchmark returns. Also the LB test constitutes 22 rejections, which corresponds to the number of rejected independence hypotheses in the summary. That way, independence is accepted for *EquAsP*, *RateUSD* and some currency series.

[37] Portmanteau proves whether a sum of squared autocorrelation coefficients differs from zero, with a χ^2_K test statistic. Intrinsically, $H_0 : \rho(1) = \rho(2) = \ldots = \rho(K) = 0$ is tested ($K \approx 2\sqrt{n}$, following a rule of thumb; $K = 20$ for monthly, $K = 90$ for daily samples).

[38] We use the BDS independence test of the software EViews, testing for different dimensions.

[39] For the diff-returns of *CredSta* we identify a significant negative first order autocorrelation. Similar correlograms are observed for the other daily credit series, too. An autoregressive (AR(1)) structure is suggested on that exposures.

[40] But non-stationary effects could entail similar results.

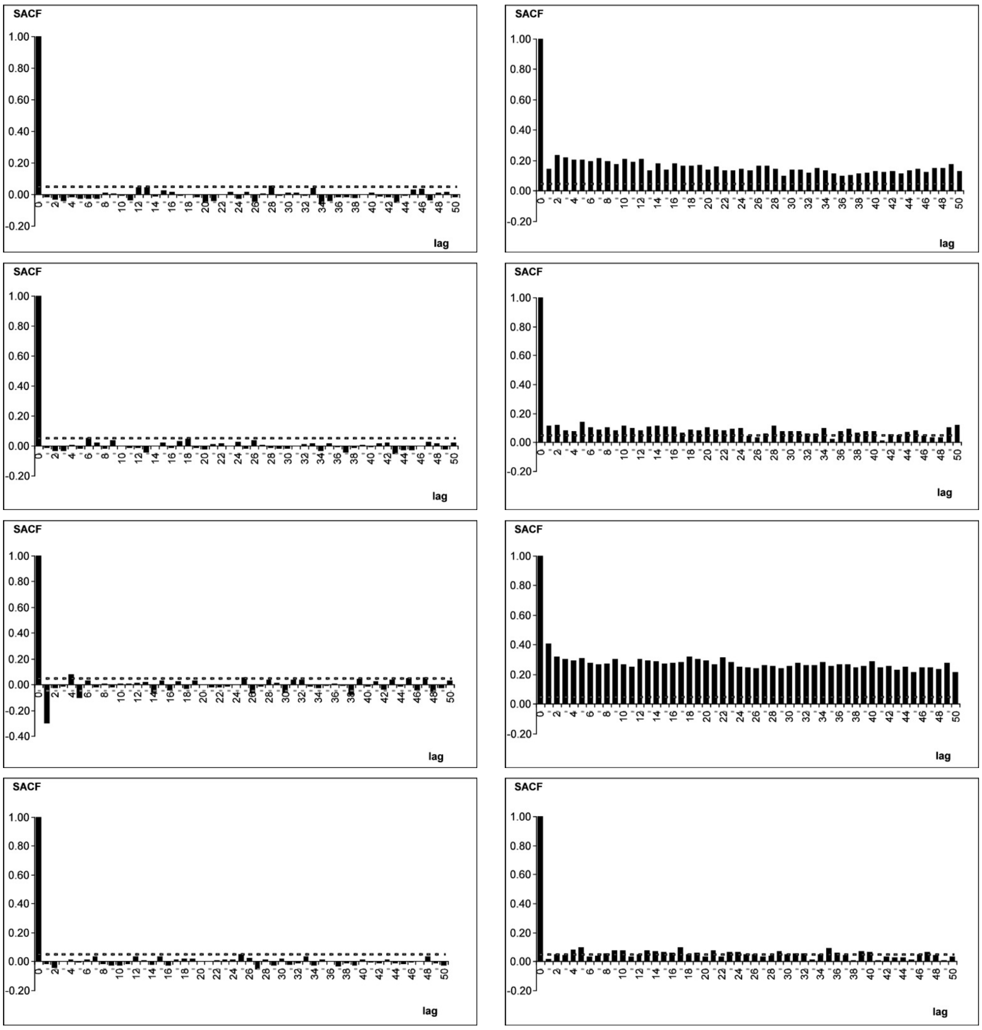


Figure 11: Correlograms for daily returns (left) and their absolutes (right) of *EquNA*, *RateUSD*, *CredSta* and *CurrUSD* (from top to bottom). The dashed lines define the approximative 95% confidence band.

number of rejected independence hypotheses in summary (standard)	22
-> thereof equities	4
-> thereof interest rates	2
-> thereof credit spreads	8
-> thereof currencies	8

Figure 12: Test summary heuristic concerning the serial independence of monthly benchmark returns.

Altogether, after having seen a surprising good independence feature on the monthly data base, the daily benchmark returns do not fulfill this hypothesis in most instances. Daily returns are established to be serially uncorrelated but dependent. A bridging to the normal distribution and the resulting correlograms gave further evidence that the daily returns cannot be Gaussian. On the other hand the testing approaches required a serially identical distribution, which was partly denied before. The interpretation of correlograms depends on the stationarity assumption, which should be queried itself and is abandoned in section 3.

Summary and conclusion on testing model assumptions:
The modelling premise of iid normally distributed benchmark returns was requested in a comprehensive survey. On an 8-year monthly data base of benchmark returns it can be barely supported, albeit for 7 of 29 indices the normal assumption was denied and serial identity was rejected 19 times. Testing independence brought only 3 rejections. This result worsens when refining the horizon to improve statistical quality: For 8-year daily samples the assumptions of normality and identity were clearly rejected, independence tests performed with only 22 rejections still best. The premise of iid normal returns fails entirely. Among the exposure classes credit spreads caused the most approximation problems, but also equities showed a very bad fit with respect to the assumptions. The use of different return types has a minor influence on the testing results as long as no modelling errors occur (for zero/ negative prices and arithmetic or log-returns). Those troubles dominated the choice of a preferred return conception, where we arrived at log-returns for equity indices and exchanges rates and at diff-returns for series of interest rates and credit spreads.

Besides the normal distribution assumption, the volatility measurement via empirical standard deviations and the transformation of volatility and VaR to other horizons are critical if returns are not iid. Due to different testing results on the monthly and daily setup the temporal closedness of the model is requested.[41] Contradictions arise when transforming the VaR on daily or monthly returns into each other or when comparing annualized volatilities. Altogether and consequently, the premises of the basic risk model have to be denied.[42] Improvements can be based on a non-stationary and heteroscedastic setup for financial returns. Since the larger problems were identified on the daily data base, we switch from this point to the daily benchmark series, working with the preferred return conception.

3 A non-stationary model for asset returns

In the previous section the demand for a non-stationary and heteroscedastic model that allows adequate (asymmetric, heavy tailed) return distributions evolved from testing the benchmark return series. Since Granger and Starica (2005) have documented the superiority of the paradigm of time-varying unconditional variance over some specifications of stationary long memory, our following approach is based on interpreting the slow decay of the SACF of absolute returns as a sign of non-stationarities in the second-moment structure (compare section 1). While ARCH-type and stochastic volatility models specialized more and more to find endogenous variance or covariance descriptions for special financial instruments, we draw the conclusion that a simple endogenous specification does not exist and assume the unconditional volatility to be exogenous to the return process. The evolution of market prices is interpreted as a manifestation of complex market conditions, driven by unknown endogenous factors. As a consequence we model the volatility deterministic, building on a non-stationary model for daily log-returns $X_t = \ln P_t - \ln P_{t-1}$ and diff-returns $X_t = P_t - P_{t-1}$ $(t = 1, \ldots, n)$ according to the modelled exposure type.[43]

We take up the conceptual framework of Herzel et al. (2005) and Drees and Starica (2002) (univariate case) for analysing the return dynamics via classical nonparametric regression with fixed equidistant design points. The vectors of financial returns are assumed to have a time-varying unconditional covariance matrix that evolves smoothly through time. The standardized residuals are modelled parametrically, allowing asymmetry and heavy tails. This leads to a multiplicative approach congruent to equation (1), with a constant mean return μ added, for a non-stationary sequence of independent random vectors $\{\mathbf{X}_t\}_t$:

$$\mathbf{X}_t = \mu + \mathbf{S}_t \varepsilon_t, \; t = 1, \ldots, n, \tag{12}$$

$$\varepsilon_1, \ldots, \varepsilon_n \text{ iid random vectors with mutually independent coordinates,}$$
$$E\varepsilon_{k,1} = 0, Var\varepsilon_{k,1} = 1, \forall k = 1, \ldots, d,$$
$$\mathbf{S}_t : [0, n] \to \mathbb{R}^{d \times d} \text{ is an invertible matrix and a smooth function of time.}$$

[41] In another case study on estimated standard deviations we surveyed the impact of the choice of periodicities and lengths of the measurement window. While caring for autocorrelation effects, the higher volatility estimate depends on the current market stadium and the influence of past extremal movements or on the fast capturing of recent (upcoming) movements, leading to a trade-off between long-memory and instantaneous changes. Problems of time transformation (via square root of time) between the measures emerged, that could be due to a lack of robustness of (small) samples or due to serial dependence. Hence, the length and granularity of financial time series is important for estimating volatility. Moreover, most volatility estimations motivated a heteroscedastic setup.

[42] A justification for the professional risk manager using similar models is, that for a multitude of exposures in a highly diversified portfolio several model shortcomings compensate each other. So arbitrary probability distributions of different exposures lead to an approximative normally distributed portfolio return, following the law of large numbers. A similar argument holds on temporal distributional changes and dependencies. Daily autocorrelations decrease on monthly returns. Backtesting procedures may confirm that effects and a holistic functionality.

[43] The model and statistical features given below work for diff-returns as well as for log-returns.

The residuals ε_t are called standardized innovations. We emphasise that this regression-type model does not exclude random effects of the volatility dynamics. The basic idea is that recent past and the next future returns depend on the same unknown exogenous economic factors, that evolve gradually through time. Those factors are included in the recent asset returns and imply the level of the unconditional (co)variance. The aim is to estimate the multivariate return dynamics only by dint of recent returns and to build up short-term forecasts of future return distributions in a similar economic environment. Three steps have to be arranged to fit the regression model to a financial time series:

1. Centering returns
The demeaned return series $\{R_t\}_t$ is defined as

$$\mathbf{R}_t = \mathbf{X}_t - \bar{\mathbf{X}}_n, \ \ t = 1, \ldots, n, \tag{13}$$

with column vectors centered componentwise by the empirical mean of the whole series, $\bar{X}_{k,n} = \frac{1}{n}\sum_{t=1}^{n} X_{k,t}$ for all $k = 1, \ldots, d$. If the estimation error of $\bar{\mathbf{X}}_t$ for μ was neglected (i.e. $\bar{\mathbf{X}}_t = \mu \implies \mathbf{R}_t = \mathbf{S}_t\varepsilon_t$), it follows that:

$$E\left(\mathbf{R}_t \mid \mathbf{R}_{t-1}, \mathbf{R}_{t-2}, \ldots\right) = E\mathbf{R}_t = \mathbf{S}_t E\varepsilon_t = 0 \tag{14}$$

$$E\left(\mathbf{R}_t\mathbf{R}_t' \mid \mathbf{R}_{t-1}, \mathbf{R}_{t-2}, \ldots\right) = E\left(\mathbf{R}_t\mathbf{R}_t'\right) = \mathbf{S}_t\mathbf{S}_t' =: \mathbf{\Sigma}^2(t) \tag{15}$$

Hence, $\{\mathbf{R}_t\mathbf{R}_t'\}_t$ is an independent sequence of matrices with pointwise expectations $\mathbf{\Sigma}^2(t)$, a smooth function of time. This offers the framework for a nonparametric regression on equidistant design points $t = \{1, \ldots, n\}$, where variances and covariances in matrices $\left\{\mathbf{\Sigma}^2(t)\right\}_t$ are estimated with standard nonparametric estimators.

2. Estimating volatilities
With the tools of classical nonparametric regression, we derive from a local polynomial regression method (local constant regression) on $\{\mathbf{R}_t\mathbf{R}_t'\}_t$, with data localized by kernel functions K_h,[44] and the method of least squares a *Nadaraya-Watson estimator*:

$$\hat{\mathbf{\Sigma}}^2(t) = \frac{\sum_{i=1}^{n} K_h(i-t)\mathbf{R}_i\mathbf{R}_i'}{\sum_{i=1}^{n} K_h(i-t)}, \tag{16}$$

with $K_h(\cdot) = \frac{1}{h}K\left(\frac{\cdot}{h}\right)$, where K is an appropriate kernel as defined in the sequel. This is the two-sided volatility estimate for the multivariate regression model (12). Regarding the included return information we have to distinguish later between the two-sided (symmetrical) and the one-sided (historical) estimation.

Herzel et al. (2005) motivate the application of nonparametric regression by theoretical results of Müller and Stadtmüller (1987) in an asymptotic context (compare section 3.2 in Herzel et al. (2005)). That way, they additionally derive propositions on confidence intervals for $\left(\Sigma_{i,j}(t)\right)_{i,j}$. We restrict ourselves later to the univariate case and outline some useful statistical results from Gürtler et al. (2009).

3. Fitting innovations
In the last step we have to model the distribution of innovations $\{\varepsilon_t\}_t$. Firstly, the innovations $\varepsilon_{k,t}$ are componentwise estimated by dint of demeaned returns $R_{k,t}$ and estimated volatilities $\hat{\sigma}_k(t) = \left(\hat{\Sigma}_{k,k}(t)\right)_{i,j}$ (the k-th diagonal element of the square root of the estimate $\hat{\mathbf{\Sigma}}^2(t)$ for $\mathbf{S}_t\mathbf{S}_t'$) in each point of time t:

$$\hat{\varepsilon}_{k,t} = \frac{R_{k,t}}{\hat{\sigma}_k(t)}, \ \ t = 1, \ldots, n. \tag{17}$$

Due to their independence it is sufficient to specify the distributions of $\hat{\varepsilon}_{k,t}$, $k = 1, \ldots, d$ univariately. The easiest approach without any extra assumption could be the empirical distribution function $\hat{F}_n^{emp}(x)$ of the series $\{\hat{\varepsilon}_t\}_t$, but this is not able to capture heavy tails.[45,46] Herzel et al. (2005) as well as Drees and Starica (2002) found the *Pearson*

[44] We use symmetric kernel functions K and K_h on compact supports $[-1, 1]$ and $[-h, h]$, respectively: $K : \mathbb{R} \to [0, \infty)$ with $\int_{-\infty}^{\infty} K(u)du = \int_{-1}^{1} K(u)du = 1$. $K_h(u) := \frac{1}{h}K\left(\frac{u}{h}\right)$, denoting K_h as rescaled kernel on a bandwidth $h > 0$.

[45] Being $x_{max} = \max_t \hat{\varepsilon}_t$ and $x_{min} = \min_t \hat{\varepsilon}_t$ of the innovation sample $\{\hat{\varepsilon}_t\}_t$, then $\hat{F}_n^{emp}(x_{max}+\delta) = 1$ for all $\delta \geq 0$ and $\hat{F}_n^{emp}(x_{min} - \epsilon) = 0$ for all $\epsilon > 0$. Consequently the probability for extreme future innovations $\varepsilon_N \notin [x_{min}, x_{max}]$, $N > n$ would equal 0. The empirical distribution function underestimates the extremes, which is an unacceptable shortcoming for risk management purposes.

[46] Moreover, the normal distribution drops out due to neglecting skewness and heavy tails.

type VII distribution to be a flexible and parsimonious family of heavy-tailed distributions.[47] It has the following one-sided density with shape parameter m and scale parameter c:

$$f_{m,c}^{VII(1)}(x) = \frac{2\Gamma(m)}{c\Gamma\left(m-\frac{1}{2}\right)\pi^{1/2}}\left(1+\left(\frac{x}{c}\right)^2\right)^{-m} I_{[0,\infty)}(x). \tag{18}$$

The one-sided Pearson VII presentation, concentrated on the positive axis, was chosen to allow for asymmetry: It is fitted separately to nonnegative innovations $\{\hat{\varepsilon}_t \mid \hat{\varepsilon}_t \geq 0\}_t$ and absolute values of negative innovations $\{-\hat{\varepsilon}_t \mid \hat{\varepsilon}_t < 0\}_t$. In the above mentioned literature the corresponding parameters (m_+, c_+) and (m_-, c_-) are estimated with maximum-likelihood methods. We solve that task with a method of moments, later on. Because of the fact that there are usually about as many positive as negative innovations in a financial time series, it may be assumed that the median of innovations is zero. Hence, the fitted one-sided Pearson VII densities $f_{m_+,c_+}^{VII(1)}$ and $f_{m_-,c_-}^{VII(1)}$ are combined as

$$f_{m_+,c_+,m_-,c_-}^{VII}(x) = \frac{1}{2}\Big(f_{m_-,c_-}^{VII(1)}(-x)I_{(-\infty,0)}(x) + f_{m_+,c_+}^{VII(1)}(x)I_{[0,\infty)}(x)\Big), \tag{19}$$

its cdf is referred as $\hat{F}^{VII}(x)$ and called *asymmetric Pearson type VII distribution* (Drees and Starica (2002)) of random innovations ε_t.[48]

Concluding, the estimated distributions $\hat{F}_{\varepsilon_k}^{VII}$ of the d independent random innovations together with (the square root of) the covariance matrix estimates $\hat{\mathbf{\Sigma}}^2(t)$ and the mean vector $\bar{\mathbf{X}}_n$ completely specify the distribution of returns $\mathbf{X}_t$ in the regression model (12).

3.1 Univariate nonparametric estimation of volatility

From now on we specialize on the univariate non-stationary model, as introduced in Drees and Starica (2002):

$$\begin{aligned} &X_t = \mu + \sigma(t)\varepsilon_t, \;\; t = 1,\dots,n, \\ &\varepsilon_1,\dots,\varepsilon_n \;\text{ iid with }\; E\varepsilon_1 = 0,\; Var\,\varepsilon_1 = 1, \\ &\sigma(t), \;\; t = 1,\dots,n, \;\text{ a smooth, deterministic function of time.} \end{aligned} \tag{20}$$

The series of log-returns $\{X_t\}_{t=1,\dots,n}$ preserves the independence assumption, but it abandons the hypothesis of serial identity, because of the unconditional time-varying volatility within the return distribution. Furthermore the asymptotic, heavy tailed modelling of the random innovations replaces the overall normal return distribution of the basic risk model.

After demeaning the return series $\{X_t\}_t$ we base the nonparametric regression on equidistant points $t = \{1,\dots,n\}$ of squared returns $\{R_t\}_t$ with the methods developed above. The Nadaraya-Watson estimator (short: NWE) for volatility estimation is obtained as:

(I) Two-sided NWE (smoother):

$$\hat{\sigma}^2(t) := \frac{\sum_{i=1}^n K_h(i-t)R_i^2}{\sum_{i=1}^n K_h(i-t)} \tag{21}$$

with $K_h(\cdot) = \frac{1}{h}K\left(\frac{\cdot}{h}\right)$ for an appropriate kernel K on a compact support $[-1,1]$.

(II) One-sided NWE (filter):

$$\hat{\sigma}_{(1)}^2(t) := \frac{\sum_{i=1}^t K_h(i-t)\widetilde{R}_i^2}{\sum_{i=1}^t K_h(i-t)} \tag{22}$$

with $K_h(\cdot) = \frac{1}{h}K\left(\frac{\cdot}{h}\right)$ and $\widetilde{R}_i := X_i - \bar{X}_{i-1}$.

The symmetric estimator (I) immediately follows from equation (16), an asymmetric version (II) that includes only past/ current returns is supplemented. The distinction between both approaches is fundamentally. The first estimation of $\sigma^2(t)$ depends on a symmetrical data base around t, using all returns R_i that are temporally close enough, i.e. all values whose design points i are within the bandwidth h around t.[49] From that perspective past and future returns are

[47] The Pearson VII family includes the Student t-distribution, the Cauchy distribution and asymptotically the Gaussian distribution. Other applications are e.g. to be seen in Kitagawa and Nagahara (1999) for standardized innovations in a stochastic volatility (state-space) model.

[48] Another approach for fitting innovations could be kernel density estimation, to stay completely within a nonparametric approach.

[49] Values outside this radius are assumed not to influence $\sigma(t)$ and are zero-weighted.

included in volatility estimation. The calculation requires a sufficient long sample and is only possible if $1 \leq t - h$ and $t + h \leq n$. Else the required band $[-h, +h]$ would go beyond the data base and boundary effects occur.[50]

While a symmetric kernel will be used to describe the dynamics of changes in a historical sample, the one-sided estimation is applied in forecasting volatility. The asymmetric estimator $\hat{\sigma}^2_{(1)}(t)$ inserts only past (centered) returns $\widetilde{R}_i$ up to time t, restricted to values within the left-sided band $[t - h, t]$. Again a sufficient long past series is required, the (unbiased) estimation is possible for points t with $1 \leq t - h$. Else a boundary effect results on the left end. Of course, in a historical sample the one-sided estimator delivers generally a bigger estimation error than its two-sided counterpart due to the lack of information.

We aim at deriving statistical results as consistency or asymptotic normality for our volatility estimates. At first we have to define an appropriate asymptotic framework, where we do not only involve an increasing number of observations but even more an increase of the frequency for observing data points on a fixed time-frame. Analytically, we rescale the observations to a unit interval and refine gradually the data base. That way, an increase of sample size n means including more observations on closer design points. Henceforth we scan the (unknown) regression function more and more precisely, until n goes to infinity.

We assume that beyond the series of returns $\{X_t\}_{t=1,\ldots,n}$ the discrete sequence $\{\sigma(t)\}_t$ is gathered from a continuous volatility function:

$$\sigma : [0, n] \longrightarrow \mathbb{R}^+_0, \; x \mapsto \sigma(x) \tag{23}$$

The aforementioned asymptotic is implemented by transforming the data series to the standardized window $[0, 1]$ with design points $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1$, that adopts the volatility values as $s\left(\frac{t}{n}\right) = \sigma(t)$ for all $t = 1, \ldots, n$. Formally we have:

$$s : [0, 1] \longrightarrow \mathbb{R}^+_0, \; y \mapsto s(y) := \sigma(ny). \tag{24}$$

This produces the transformed multiplicative return model:

$$\begin{aligned} X_{t,n} \;\; &= \;\; \mu + s\left(\frac{t}{n}\right)\varepsilon_{t,n}, \;\; t = 1, \ldots, n, \\ &\varepsilon_{1,n}, \ldots, \varepsilon_{n,n} \;\text{ iid with }\; E\varepsilon_{1,n} = 0, \; Var\,\varepsilon_{1,n} = 1, \\ &s\left(\frac{t}{n}\right), \;\; t = 1, \ldots, n, \;\; \text{a smooth, deterministic function of time.} \end{aligned} \tag{25}$$

The corresponding nonparametric estimators obviously are:

(I) Two-sided transformed NWE:

$$\hat{s}^2_{h_n}(u) := \frac{\sum_{i=1}^{n} K_{h_n}\left(\frac{i}{n} - u\right) R^2_{i,n}}{\sum_{i=1}^{n} K_{h_n}\left(\frac{i}{n} - u\right)}, \tag{26}$$

with $u \in [0, 1]$, $h_n := h(n)$ and $K_{h_n}(\cdot) = \frac{1}{h_n} K\left(\frac{\cdot}{h_n}\right)$.

(II) One-sided transformed NWE:

$$\hat{s}^2_{(1)h_n}(u) := \frac{\sum_{i=1}^{\lfloor un \rfloor} K_{h_n}\left(\frac{i}{n} - u\right) \widetilde{R}^2_{i,n}}{\sum_{i=1}^{\lfloor un \rfloor} K_{h_n}\left(\frac{i}{n} - u\right)}, \tag{27}$$

with $u \in [0, 1]$, $h_n := h(n)$, $K_{h_n}(\cdot) = \frac{1}{h_n} K\left(\frac{\cdot}{h_n}\right)$ and $\widetilde{R}_{i,n} = X_{i,n} - \bar{X}_{i-1}$.

For proving consistency and asymptotic normality of the estimators $\hat{s}^2_{h_n}(u)$ and $\hat{s}^2_{(1)h_n}(u)$ certain conditions on the kernel K, the bandwidth h_n and the smoothness of the (transformed) volatility function $s(\cdot)$ are required:

[50] Several approaches for treating the boundary region $t \in [0, h)$ and $t \in (n - h, n]$ exist in the literature of nonparametric curve estimation. Fan and Yao (2003) list e.g. special *boundary kernels*, methods of *reflection* and *transformation* or local polynomial fitting of a higher degree. In general, the order of magnitude of the bias is different in the interior and near the boundaries. This is to be seen in the subsequent analysis as the optimal two-sided (interior) bandwidth is of order $n^{4/5}$ while the optimal bandwidth of the left-sided estimator has size $n^{2/3}$, that could be interpreted as a boundary corrected estimator for the right interval boundary.

(C1) Let $K : \mathbb{R} \to [0, \infty)$ be a symmetrical density with compact support $[-1, 1]$, i.e. the kernel has the features: (i) $K(v) = 0 \; \forall v \notin [-1, 1]$, (ii) $\int_{-\infty}^{\infty} K(v)dv = 1$, (iii) $\int_{-\infty}^{\infty} vK(v)dv = 0$.

(C2) Let K be continuous with a limited first derivation K'.

(C3) $K_{h_n}(\cdot) = \frac{1}{h_n} K\left(\frac{\cdot}{h_n}\right)$ with restrictions to the bandwidth h_n: (i) $h_n \overset{n\to\infty}{\longrightarrow} 0$, (ii) $nh_n, \ldots, nh_n^4 \overset{n\to\infty}{\longrightarrow} \infty$, $nh_n^6, nh_n^7, \ldots \overset{n\to\infty}{\longrightarrow} 0$, (iii) $nh_n^5 \overset{n\to\infty}{\longrightarrow} C^2 \geq 0$.

(C4) Let s^2 be two times continuous differentiable.

(C5) Let rvs $\varepsilon_{1,n}, \ldots, \varepsilon_{n,n}$ be iid with $E\varepsilon_{1,n} = 0$, $Var\, \varepsilon_{1,n} = 1$ and $E\left|\varepsilon_{1,n}\right|^{4+\delta} < \infty$ for a $\delta > 0$ and $n \in \mathbb{N}$.

Two-sided volatility estimation

A minimum requirement to a developed, feasible estimator is consistency, that we prove first for the both-sided estimate $\hat{s}_{h_n}(u)$ of the volatility function $s(u)$, $u \in [0, 1]$. Heuristically, an increase of the sample size should imply that the estimator converges to the parameter to be estimated. That way, a good estimate has to be asymptotically unbiased and its variance should converge to zero as n goes to infinity. Finally, the stochastical convergence is to be concluded, as presented in proposition 2.1 of Gürtler et al. (2009):

(P1) *Under the conditions (C1) - (C5) in setup (25) the sequence* $\left(\hat{s}^2_{h_n}(u)\right)_{n\in\mathbb{N}}$ *of estimators for* $s^2(u)$ *is consistent for all* $u \in (0, 1)$.

Furthermore they extended the consistency result by inspecting the rate of convergence. $\sqrt{nh_n}\left(\hat{s}^2_{h_n}(u) - s^2(u)\right)$ is limited (in probability) due to an asymptotical bias, a finite variance and an asymptotic normal distribution. The convergence in distribution was proved with the central limit theorem of Lindeberg-Feller. Hence, it holds that (proposition 2.2, Gürtler et al. (2009)):

(P2) *Let* $C \geq 0$ *and* $V := E\varepsilon^4_{1,n} - 1 \in (0, \infty)$ *and the conditions (C1) - (C5) be satisfied. Then the sequence of estimators* $\left(\hat{s}^2_{h_n}(u)\right)_{n\in\mathbb{N}}$ *for* $s^2(u)$ *is asymptotic normally distributed for all* $u \in (0, 1)$ *in terms of*

$$\sqrt{nh_n}\left(\hat{s}^2_{h_n}(u) - s^2(u)\right) \overset{D}{\longrightarrow} \mathcal{N}\left(\beta(u), \tau^2(u)\right), \quad \textit{where} \tag{28}$$
$$\beta(u) = \frac{C}{2}\left(s^2(u)\right)'' \int_{-1}^{1} v^2 K(v)dv,$$
$$\tau^2(u) = V s^4(u) \int_{-1}^{1} K^2(v)dv.$$

The asymptotic normality result of the two-sided variance estimate reveals some more insights: For the finite approximation of $s^2(u)$ by $\hat{s}^2_{h_n}(u)$ the slowest error terms have a $\sqrt{nh_n}$ rate of convergence. Concluding from the quoted proposition, for a sufficiently large n the pointwise approximation is nearly distributed as:

$$\hat{s}^2_{h_n}(u) - s^2(u) \approx \mathcal{N}\left(\frac{h_n^2}{2}\left(s^2(u)\right)'' \int_{-1}^{1} v^2 K(v)dv, \frac{V}{nh_n} s^4(u) \int_{-1}^{1} K^2(v)dv\right). \tag{29}$$

Here, the approximate bias has a negligible magnitude relative to the variance term. Thus, an approximative confidence interval for $s^2(u)$ can be simplistic implemented with Gaussian quantiles, built on a normal distribution centered at $\hat{s}^2_{h_n}(u)$ and a variance as above. Moreover, discussions on an optimal bandwidth continue on that approximate normal parameters. Since the mean squared error is $MSE\hat{s}^2_{h_n}(u) = Bias^2\hat{s}^2_{h_n}(u) + Var\hat{s}^2_{h_n}(u)$, minimizing the function with respect to the bandwidth will provide the optimal trade-off between bias and variance. Our later considerations regarding optimal bandwidths extent that result to a MISE criterion.

One-sided volatility estimation

A similar statistical analysis was executed for the historical, left-sided volatility estimator $\hat{s}^2_{(1)h_n}(u)$. Concluding, the statistical results of the two-sided estimate can be transfused to the one-sided counterpart, conditioned on a faster convergence rate of the bandwidth h_n. For the one-sided NWE condition (C3) has to be replaced with (C3') while the other premises (C1), (C2), (C4), (C5) persist:

(C3') $K_{h_n}(\cdot) = \frac{1}{h_n} K\left(\frac{\cdot}{h_n}\right)$ with restrictions to the bandwidth h_n: (i) $h_n \overset{n\to\infty}{\longrightarrow} 0$, (ii) $nh_n, nh_n^2 \overset{n\to\infty}{\longrightarrow} \infty$, $nh_n^4, nh_n^5, \ldots \overset{n\to\infty}{\longrightarrow} 0$, $nh_n^3 \overset{n\to\infty}{\longrightarrow} D^2 \geq 0$.

The one-sided-estimator $\hat{s}^2_{(1)h_n}(u)$ is asymptotically unbiased, its variance tends to zero for large samples and the stochastical convergence follows. Gürtler et al. (2009) provide with proposition 2.3.:

(P3) *Under the conditions (C1), (C2), (C3'), (C4), (C5) in setup (25) the sequence* $\left(\hat{s}^2_{(1)h_n}(u)\right)_{n\in\mathbb{N}}$ *of estimators for* $s^2(u)$ *is consistent for all* $u \in (0,1]$.

The consistency result for $\hat{s}^2_{(1)h_n}(u)$ is not only valid at interior points of $[0,1]$ (as it was the case for the two-sided approach), but also at the right frontier. This makes the estimator consistently applicable for forecasting volatility and return distributions, respectively. Again $\sqrt{nh_n}\left(\hat{s}^2_{(1)h_n}(u) - s^2(u)\right)$ is limited (in probability) due to an asymptotical bias, a finite variance and a convergence in distribution to a Gaussian. Hence, it holds that (proposition 2.4, Gürtler et al. (2009)):

(P4) *Let* $D \geq 0$ *and* $V := E\varepsilon^4_{1,n} - 1 \in (0,\infty)$ *and the conditions (C1), (C2), (C3'), (C4), (C5) be satisfied. Then the sequence of estimators* $\left(\hat{s}^2_{(1)h_n}(u)\right)_{n\in\mathbb{N}}$ *for* $s^2(u)$ *is asymptotic normally distributed for all* $u \in (0,1]$ *in terms of*

$$\sqrt{nh_n}\left(\hat{s}^2_{(1)h_n}(u) - s^2(u)\right) \overset{D}{\longrightarrow} \mathcal{N}\left(\beta_{(1)}(u), \tau^2_{(1)}(u)\right), \ \textit{where} \tag{30}$$
$$\beta_{(1)}(u) = 2\,D\left(s^2(u)\right)' \int_{-1}^{0} vK(v)dv,$$
$$\tau^2_{(1)}(u) = 4\,Vs^4(u)\int_{-1}^{0} K^2(v)dv.$$

It is possible to draw similar conclusions from the asymptotic normality of $\hat{s}^2_{(1)h_n}(u)$, as done for the two-sided NWE. It follows from the last proposition for sufficiently large n pointwise:

$$\hat{s}^2_{(1)h_n}(u) - s^2(u) \approx \mathcal{N}\left(2h_n\left(s^2(u)\right)' \int_{-1}^{0} vK(v)dv, \frac{4\,V}{nh_n}s^4(u)\int_{-1}^{0} K^2(v)dv\right). \tag{31}$$

With that result, an approximative confidence interval for $s^2(u)$ can be simplistic implemented in terms of normal quantiles, centered at $\hat{s}^2_{(1)h_n}(u)$ and scaled by the standard deviation from above. Moreover, discussions on optimal one-sided bandwidths, that minimize the MSE or MISE of $\hat{s}^2_{(1)h_n}(u)$, continue in the subsequent section.

Choice of kernel and bandwidth

In the prevalent literature it is established, that the choice of the kernel function plays a relatively unimportant role compared with the optimal bandwidth for nonparametric regression. Although different kernels perform very similar in large samples, we have to remember some of the smoothness conditions,[51] which exclude standards as rectangle-, triangular- or normal kernels. We recommend a polynomial of fourth degree, also called *biweight kernel*:[52]

$$K(u) := \begin{cases} \frac{15}{16}(1-u^2)^2 & , \quad |u| \leq 1 \\ 0 & , \quad else \end{cases} . \tag{32}$$

Using the biweight kernel in the sequel, we turn to the task of bandwidth selection. The bandwidth is also called smoothing parameter because one has to find a trade-off between over- and undersmoothing.[53] Discussions on optimal

[51] Amongst others we need a continuous differentiable kernel with a compact support.

[52] Following Fan and Yao (2003), this kernel is from the *'symmetric Beta family'* $K_\gamma(u) = \frac{1}{\mathcal{B}\left(\frac{1}{2}, \gamma+1\right)}\left(1-u^2\right)^\gamma I_{[-1,1]}(u)$ with beta-integral $\mathcal{B}(\alpha_1,\alpha_2) = \int_0^1 (1-y)^{\alpha_1-1} y^{\alpha_2-1} dy$ as the special case $\gamma = 2$.

[53] Oversmoothing means in terms of NWEs to build an average over a too large neighbourhood of return points (large bandwidth), where recent return information is dominated; a very smooth shape of the regression function results (small variance, but biased). Undersmoothing averages over a very small neighbourhood (small bandwidth), where only a few recent data points are included; a rough shape (small bias, but large variance) of the estimated volatility is the consequence.

bandwidths require an error measure. Local bandwidth optimization can be based on minimizing the MSE with respect to h_n for a point in time. Doing so, the afore derived asymptotic results are constraining. An average of MSE optimal bandwidths over all design points corresponds to the MASE criterion. Superior to this is to integrate the squared error of the estimator $\hat{s}(\cdot)$ to the volatility function $s(\cdot)$ over the (standardized) horizon and to minimize that functional. We arrive at the MISE and get for sufficient large n the following global optimal bandwidths (compare Gürtler et al. (2009)):

(I) For the two-sided (transformed) NWE (26):

$$h_n^{opt} = n^{-\frac{1}{5}} \left(\frac{V\left(\int_0^1 s^4(u)du\right)\left(\int_{-1}^1 K^2(v)dv\right)}{\left(\int_0^1 \left((s^2(u))''\right)^2 du\right)\left(\int_{-1}^1 v^2 K(v)dv\right)^2} \right)^{\frac{1}{5}} . \tag{33}$$

(II) For the one-sided (transformed) NWE (27):

$$h_{(1)n}^{opt} = n^{-\frac{1}{3}} \left(\frac{V\left(\int_0^1 s^4(u)du\right)\left(\int_{-1}^0 K^2(v)dv\right)}{2\left(\int_0^1 \left((s^2(u))'\right)^2 du\right)\left(\int_{-1}^0 vK(v)dv\right)^2} \right)^{\frac{1}{3}} . \tag{34}$$

Those criteria can be used directly only for simulation studies (as in section 4.1), where the volatility function $s(\cdot)$ is a predefined input. For empirical samples it is naturally the task to estimate $s(\cdot)$, so bandwidth criteria that are based on that function or its derivatives are problematical.[54] One usual way out is the *cross-validation method* (CV), that determines the optimal smoothing parameter solely with the return series and without any knowledge about the regressed volatility function. For empirical samples $\{X_t\}_{t=1,\ldots,n}$ it is not necessary to transform the setup first for estimating $s\left(\frac{t}{n}\right)$, rather the volatility $\sigma(t)$ should be estimated directly (as done in section 4.2). Thus, we turn back the transformation, $h = nh_n$, and motivate the CV-criterion in the original regression model (20).

The basic idea is the 'leave-one-out prediction' over the discrete design. In our context, the variance $\sigma^2(\cdot)$ has to be reestimated for each point $j = 1, \ldots, n$, without using the actual observation R_j^2 (or $\widetilde{R}_j^2$, respectively) itself. We modify the nonparametric volatility estimators (21) and (22) to accordant cross-validation estimators (CVEs) $\hat{\sigma}_h^{2(j)}(j)$ and $\hat{\sigma}_{(1)h}^{2(j)}(j)$. The bandwidth selection criterion is to minimize the CV-function, which is the sum of squared differences between returns R_j^2 and CVEs $\hat{\sigma}_h^{2(j)}(j)$ (alternatively $\widetilde{R}_j^2$ and $\hat{\sigma}_{(1)h}^{2(j)}(j)$).

(I) The two-sided CV-function and CVE (untransformed) are:

$$CV(h) = \frac{1}{n}\sum_{j=1}^{n}\left(R_j^2 - \hat{\sigma}_h^{2(j)}(j)\right)^2 = \frac{1}{n}\sum_{j=1}^{n}\left(\frac{\sum_{i=1}^{n} K_h(i-j)\left(R_j^2 - R_i^2\right)}{\sum_{i=1,i\neq j}^{n} K_h(i-j)}\right)^2 \overset{!}{\longrightarrow} \min_{h>1}$$
$$\text{where } \hat{\sigma}_h^{2(j)}(j) = \frac{\sum_{i=1,i\neq j}^{n} K_h(i-j)R_i^2}{\sum_{i=1,i\neq j}^{n} K_h(i-j)}, \quad h > 1 \quad (\forall j = 1, \ldots, n) \tag{35}$$

(II) The one-sided CV-function and CVE (untransformed) are:

$$CV_{(1)}(h) = \frac{1}{n-1}\sum_{j=2}^{n}\left(\widetilde{R}_j^2 - \hat{\sigma}_{(1)h}^{2(j)}(j)\right)^2 = \frac{1}{n-1}\sum_{j=2}^{n}\left(\frac{\sum_{i=1}^{j-1} K_h(i-j)\left(\widetilde{R}_j^2 - \widetilde{R}_i^2\right)}{\sum_{i=1}^{j-1} K_h(i-j)}\right)^2 \overset{!}{\longrightarrow} \min_{h>1}$$
$$\text{where } \hat{\sigma}_{(1)h}^{2(j)}(j) = \frac{\sum_{i=1}^{j-1} K_h(i-j)\widetilde{R}_i^2}{\sum_{i=1}^{j-1} K_h(i-j)}, \quad h > 1 \quad (\forall j = 2, \ldots, n) \tag{36}$$

The CV-optimal bandwidths h^{CV} and $h^{CV_{(1)}}$ for the two- and one-sided volatility estimation are numerically found via a value table for integers $h \geq 2$ or via analysing the resulting CV-plot.

[54] So called *plug-in* methods develop kernel estimators for the unknown volatility $s(\cdot)$ and its derivatives $s^2(\cdot)''$ or $s^2(\cdot)'$ and plug them into the above bandwidth formula, with an iterative procedure estimating optimal bandwidths, compare Gasser et al. (1991).

3.2 Fitting innovations

In the sequel we shortly introduce the theory for fitting the Pearson type VII distribution to a series of estimated innovations $\{\hat{\varepsilon}_t\}_{t=1,\ldots,n}$ with a method of moments. Because the residual sample is evaluated in each point of time t by the ratio of a demeaned return and an estimated volatility from the regression-type model (20), we have to distinct between innovation estimates on a two-sided and one-sided data base:

(I) The two-sided innovation estimators follow as:

$$\hat{\varepsilon}_t = \frac{R_t}{\hat{\sigma}(t)}, \ \ t = 1, \ldots, n \tag{37}$$

(II) Accordingly, the one-sided innovation estimators are:

$$\hat{\varepsilon}_t = \frac{\widetilde{R}_t}{\hat{\sigma}_{(1)}(t)}, \ \ t = 1, \ldots, n \tag{38}$$

Agreeing with Drees and Starica (2002), Herzel et al. (2005), Mikosch and Starica (2004a) or Kitagawa and Nagahara (1999), the Pearson VII distribution can capture some heavy tailed innovations quite nicely.[55] The goal is to fit the one-sided Pearson VII density $f_{m,c}^{VII(1)}$ from (18) separately to nonnegative and absolutes of negative realisations of ε_t. Under the assumption that the median is close to zero, both one-sided densities are combined to the whole asymmetric presentation (19).

The Pearson VII density is part of the Pearson systems, which can be studied in Johnson and Kotz (1970).[56] The *symmetric Pearson type VII density*[57] $f_{m,c}^{VII(2)}$ satisfies the equation

$$p(x) = k\left(d_0 + d_2 x^2\right)^{-(2d_2)^{-1}}, \tag{39}$$

where $d_0 > 0$, $d_2 > 0$ and k chosen to satisfy $\int_{-\infty}^{+\infty} p(x)dx = 1$. Johnson and Kotz (1970) estimate the parameters of the Pearson system via a method of moments, especially:

$$d_0 = (4\beta_2 - 3\beta_1)(10\beta_2 - 12\beta_1 - 18)^{-1}\mu_2, \tag{40}$$

$$d_2 = (2\beta_2 - 3\beta_1 - 6)(10\beta_2 - 12\beta_1 - 18)^{-1}, \tag{41}$$

where $\mu_2(X) = EX^2$, $\beta_1(X) = \left(\frac{E(X-EX)^3}{(E(X-EX)^2)^{3/2}}\right)^2$ (squared skewness) and $\beta_2(X) = \frac{E(X-EX)^4}{(E(X-EX)^2)^2}$ (kurtosis). Since the Pearson VII density $p(x)$ is symmetric, $\beta_1 = 0$, and $\beta_2 > 3$ is required due to a positive d_2. This means, that the conception is only applicable for samples that are heavier than a normal distribution.

Gürtler et al. (2009) provide a link from (39) to the initially used expression (18) and achieve moment estimators for m and c of the one-sided Pearson VII density $f_{m,c}^{VII(1)}(x)$. Those are:

$$m = \frac{1}{2d_2} = \frac{5\beta_2 - 9}{2\beta_2 - 6}, \tag{42}$$

$$c = \sqrt{2m\, d_0} = \sqrt{\frac{2\beta_2\mu_2}{\beta_2 - 3}}. \tag{43}$$

By inserting the empirical moment estimators for β_2 and μ_2 of subsamples $\varepsilon_+ := \{\hat{\varepsilon}_t \mid \hat{\varepsilon}_t \geq 0\}_t$ and $\varepsilon_- := \{-\hat{\varepsilon}_t \mid \hat{\varepsilon}_t < 0\}_t$ we fit $f_{m,c}^{VII(1)}$ to gains and to losses. Therewith we obtain the parsimonious parametrization m_+, c_+,

[55] This will also be demonstrated in our empirical studies in section 4.2.

[56] Every member of the system has a probability density function $p(x)$ that solves a differential equation of form $\frac{1}{p}\frac{dp}{dx} = -\frac{a+x}{d_0+d_1x+d_2x^2}$, where a, d_0, d_1 and d_2 are real shape parameters.

[57] The symmetric expression is defined on the whole real axis $x \in \mathbb{R}$ and is connected to the one-sided equivalent via $f_{m,c}^{VII(1)}(x) = 2 \cdot f_{m,c}^{VII(2)}(x)\, I_{[0,\infty)}(x)$.

m_-, c_- for the asymmetric Pearson VII density (19).[58] The estimation of β_2 can be reduced to:

$$\hat{\beta}_2(\varepsilon_+) = \frac{E\varepsilon_+^4}{E\varepsilon_+^2}\,,\ \hat{\beta}_2(\varepsilon_-) = \frac{E\varepsilon_-^4}{E\varepsilon_-^2} \tag{44}$$

Moreover, the symmetric Pearson type VII distribution equates to a scaled Student-t distribution with $2m-1$ df.[59] We get

$$f_{m,c}^{VII(2)}(x) = \frac{1}{\gamma}\, g_{2m-1}\left(\frac{x}{\gamma}\right),\ \text{with } \gamma := \frac{c}{\sqrt{2m-1}} \tag{45}$$

$$\text{and } F_{m,c}^{VII(2)}(x) = \int_{-\infty}^{x} \frac{1}{\gamma}\, g_{2m-1}\left(\frac{y}{\gamma}\right) dy \overset{z:=y/\gamma}{=} \int_{-\infty}^{x/\gamma} \frac{1}{\gamma}\, g_{2m-1}(z)\,\gamma\, dz = G_{2m-1}\left(\frac{x}{\gamma}\right), \tag{46}$$

where $g_.$ and $G_.$ are the density and cdf of a Student-t rv. Because of the simple transformation from the Pearson type VII to the Student-t distribution, whose quantiles are looked up in tables, we developed the idea to implement the non-stationary model (20) for the task of VaR calculation as a factor-based model. The univariate Value at Risk $VaR^{1-\alpha}(t)$ of an exposure $w(t)$ at time t can be modelled with a yield X_t following the regression model, as the product of $w(t)$ with a nonparametric estimated volatility $\hat{\sigma}_{(\cdot)}(t)$ and the α-quantile of a Pearson VII innovation. With $u_{m,c;\alpha}$ and $t_{2m-1;\alpha}$ being the corresponding α-quantiles it immediately follows:

$$t_{2m-1;\alpha} = \frac{u_{m,c;\alpha}}{\gamma} \iff u_{m,c;\alpha} = \gamma\, t_{2m-1;\alpha}.$$

Regarding the one-sided presentation, some attention has to be paid to

$$F_{m,c}^{VII(1)}(x) = 2 \cdot F_{m,c}^{VII(2)}(x) \implies \alpha^{(1)} = 2\alpha \ \text{ (e.g. for } \alpha = 1\%).$$

Finally, from the Student-t connection follows a restriction to the shape parameter m, that ensures the asymptotic normality of the NWEs $\hat{s}^2_{h_n}(u)$ and $\hat{s}^2_{h_n(1)}(u)$, respectively. Amongst others, (C5) $E\left|\varepsilon_{1,n}\right|^{4+\delta} < \infty$ ($\delta > 0$) was required in the asymptotic theory. As all (central) moments μ_k with $n > k$ exist for a t_n-distributed rv, the Pearson VII fit of innovations has to be conform with

$$2m - 1 > 4 \iff m > \frac{5}{2}.$$

If the set of innovations was symmetric (special case, not observed in general), the assumption $Var\,\varepsilon_1 = 1$ for iid $\{\varepsilon_t\}_t$ implies additionally the relation $c = \sqrt{2m-3}$ of shape parameters.

4 Simulation and empirical studies

So far we have provided the necessary statistical tools for estimating the volatility structure $(\sigma(t))_t$ nonparametrically and for approximating estimated innovations $\{\hat{\varepsilon}_t\}_t$ by the asymmetric Pearson type VII distribution via a method of moments. These are the main steps for fitting the regression model to a financial time series. In the first part of this section we apply the theory to simulations, that challenge the asymptotical results for finite samples by predefining a volatility function or a price-process, perturbated by a Pearson VII noise, and reestimating. In the second part, the whole regression model is applied for empirical studies on the benchmark indices, introduced in section 2, and evaluated in a backtesting framework.

[58]Special attention should be paid to that method of moments in the one-sided implementation: For estimating $\hat{\beta}_2(\varepsilon_+)$ we recommend to include the hypothetical counterpart $-\varepsilon_+$ of the one-sided set and to work on the symmetrical base $\{\varepsilon_+, -\varepsilon_+\}$ with standard formula. Else $E\varepsilon_+ \neq 0$ and $\hat{\beta}_2$ builds powers on a falsely centered rv. Proceed analogically for $\hat{\beta}_2(\varepsilon_-)$.

[59]Student's df is a measure for the heavy-tailedness and is called *tail index point estimate* in the common literature, e.g. Drees and Starica (2002).

4.1 Simulation experiment

A) Firstly, we predefine a heteroscedastic volatility function (standard deviation) as:

$$\sigma^a(t) = \frac{1}{10}\left(\sin\left(\frac{2\pi}{100}t\right)+1\right) \quad t\in[0,500], \tag{47}$$

$$s^a(u) = \frac{1}{10}\left(\sin(10\pi u)+1\right) \quad u\in[0,1]. \tag{48}$$

This corresponds to a multimodal oscillation with 5 periods in the interval $[0, 500]$. The values on the discrete design $1, \ldots, 500$ could be thought as annualized volatilities $\sigma^a(t)$ at the end of days t, observed over two years (250 trading days p.a.). The transformed version is $s^a(u)$. The example focusses first on reestimating the volatility in a simplified case of the standardized regression model (25), with a return expectation $\mu = 0$ and innovations distributed as $\varepsilon_{1,n} \sim \mathcal{N}(0, 1)$.[60] The main point is to estimate the heteroscedastic part of the simulated series via the two-sided $\hat{s}^2_{h_n}(u)$ and the one-sided NWE $\hat{s}^2_{(1)h_n}(u)$ and to observe the influence of sample size to the estimator's fit.[61] Bandwidths are optimized with the MISE-criterion.

As discussed in section 3.1 the asymptotic works by increasing the data density on a fixed time-frame $[0, 1]$. A larger sample size n enables more and closer observations, the predefined function is scanned more precisely. Figure 13 displays the nonparametric curve estimation of the volatility function s^a for different sample sizes in a median simulation of 65 repeats.[62] We discover visually that the fit improves with the sample size for both NWEs. The median SSE increases considerably slower than the sample size, for very large samples it decreases absolutely. The approximation of the predefined volatility $s^a(u)$ by nonparametric estimators $\hat{s}_{h_n}(u)$ and $\hat{s}_{(1)h_n}(u)$ is already noticeable for smaller samples of 100 or 500 points and quite satisfying using 1000 design points. Due to its additional future information, the two-sided estimator leads generally to a better and smoother fit. The left-sided approach lags behind, since the volatility does not increase until the first extremal event happens, and after a series of shocks it decays typically slower. We did not correct boundary effects, implying the first nh_n points to be distorted, for the two-sided estimator also the last nh_n values. On the right boundary both estimators are nearly the same (neglecting different bandwidths). For the sample size of 5000 both estimates cover the predetermined function visually excellent.

To appreciate the goodness of cross-validation, we estimate optimal bandwidths in the simulation example again with the CV method. Although we do not include any knowledge of s^a, the outcomes are quite similar parameters and volatility graphs. For the 5000-point setup the resulting bandwidths are 0.030 (two-sided) and 0.015 (one-sided), that are very close to the MISE-optimal smoothing parameters (compare figure 13), with absolute differences lower/ equal 0.0025, which may also be caused by the chosen grid pattern of widths 0.005. The SSEs (0.54153 for two-sided, 1.57060 for one-sided setup) are marginally larger, but the volatility estimators $\hat{s}_{h_n}(u)$ or $\hat{s}_{(1)h_n}(u)$ deliver again an excellent fit to the predetermined $s^a(u)$, that is visually very similar to the bottom graph of figure 13.

B) In the next step we expand the simulation to a simple, discrete price process $\{P_t\}_{t=0,\ldots,500}$, where $\sigma^a(t)$ on $[0, 500]$ is involved as time-variant part of an annualized volatility function $\widetilde{\sigma}^a(t)$, that additionally includes a time-invariant component σ_0. Moreover, a constant trend μ is modelled and a heavy-tailed approach for the innovations is chosen:

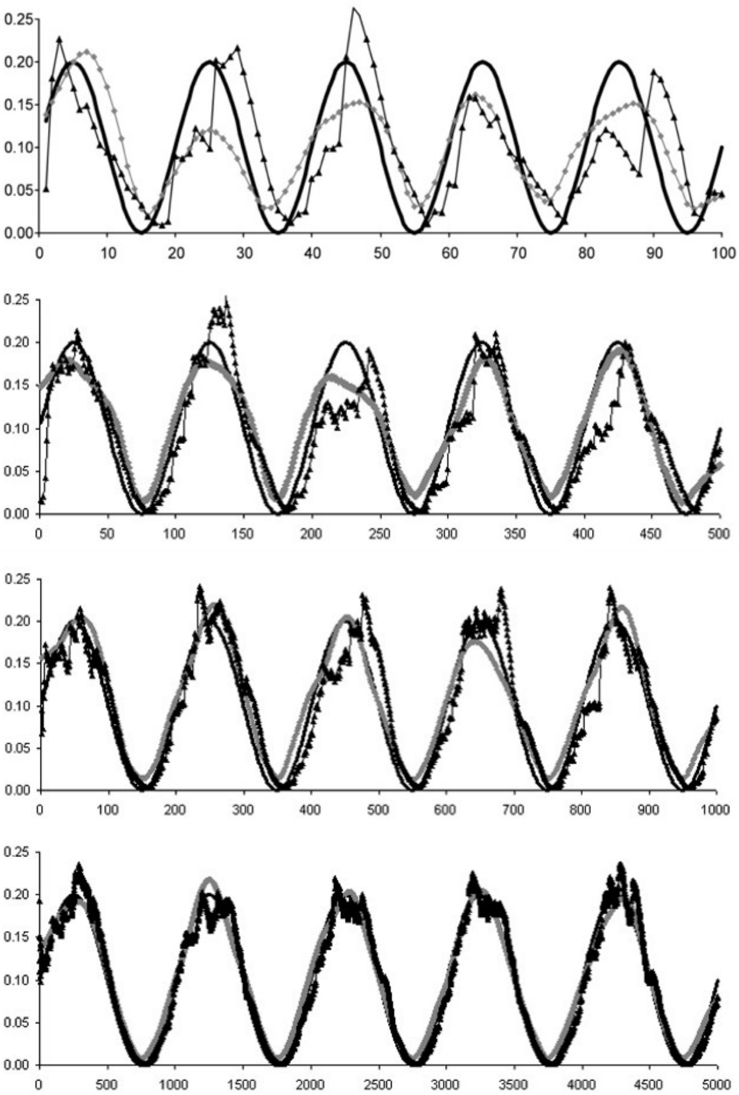
$$\begin{aligned} P_t &= P_{t-1}\, e^{X_t}\ ,\ P_0 := 1000\ ,\ t=1,\ldots,500, \quad \text{where} \\ X_t &= \mu + \sigma(t)\varepsilon_t, \text{ with } \mu := \frac{15\%}{250}, \\ \sigma(t) &= \frac{\widetilde{\sigma}^a(t)}{\sqrt{250}},\ \widetilde{\sigma}^a(t) := \sigma^a(t)+\sigma_0 \text{ and } \sigma_0 := 10\%, \\ \varepsilon_t &\sim Pearson^{VII(2)}_{m,c} \text{ with } m := 4 \text{ and } c := \sqrt{5}. \end{aligned} \tag{49}$$

The log-return of prices, $X_t = \ln(P_t/P_{t-1}) = \ln e^{X_t} = \mu + \sigma(t)\varepsilon_t$ corresponds to the (untransformed) regression approach (20). Innovations ε_t are modelled as symmetrical Pearson VII distributed rvs with equal shape parameters m, c

[60] We include later the proposed innovation modelling. The alternative choice of an expectation $\mu = const$ has a negligible influence on the simulation example due to centering the returns first of all. Assuming $\mu = 0$ is just for the sake of convenience.

[61] The simulation was implemented in a C programme, using the Box-Muller method for transforming uniform to normal random numbers. In addition we wrote VBA-code for MS Excel, to be used for smaller samples.

[62] By calculating a sum of squared errors $SSE\ \hat{s}_{h_n} := \sum_{i=1}^{n}\left(\hat{s}_{h_n}\left(\frac{i}{n}\right) - s\left(\frac{i}{n}\right)\right)^2$ of the estimator relative to the predefined volatility function for each simulated sample and ordering paths by the extent of their SSE we select the median simulation.



	BW two-sided	error (SSE)	BW one-sided	error (SSE)
100 points	0.06013	0.15739	0.06377	0.30549
500 points	0.04358	0.25614	0.03730	0.59948
1000 points	0.03794	0.28939	0.02960	0.75579
5000 points	0.02750	0.46056	0.01731	1.36289

Figure 13: Median simulation (65 samples ordered by SSE) of volatility curve estimation on s^a (continuous line, black) for 100, 500, 1000 and 5000 equidistant design points (from top to bottom) by two-sided NWEs (rhombuses, grey) and one-sided NWEs (triangles, black) with MISE-optimal bandwidths (BW).

for gains and losses. Figure 14 displays a representative simulated path of prices[63] and the nonparametric estimated volatilities relative to the original volatility in the course of time.

Here, the fit is quite acceptable despite the small sample size of 500 and the strong influence of random heavy-tailed innovations ε_t. The comparison to the price graph shows, that the nonparametric estimators are able to capture fast phases of market shocks and increased volatility, respectively.[64]

[63] Innovations were simulated in MS Excel via uniform$[0, 1]$ random numbers α, that where interpreted as probabilities corresponding to the α-quantile of a Student-t_{2m-1} rv. The Pearson VII random number is then $u_{m,c;\alpha} = \frac{c}{\sqrt{2m-1}} t_{2m-1;\alpha}$.

[64] Repeats of the simulation may produce very different paths, because the volatility level (medium 20% p.a.) dominates the price process more than its trend ($\mu = 15\%$ p.a.) and the incident of extreme innovations may cause considerable shocks. Due to the extent of random innovations, the NWEs may have different optimal bandwidths and a different developing of estimates adapted to the realized volatility of the simulated sample.

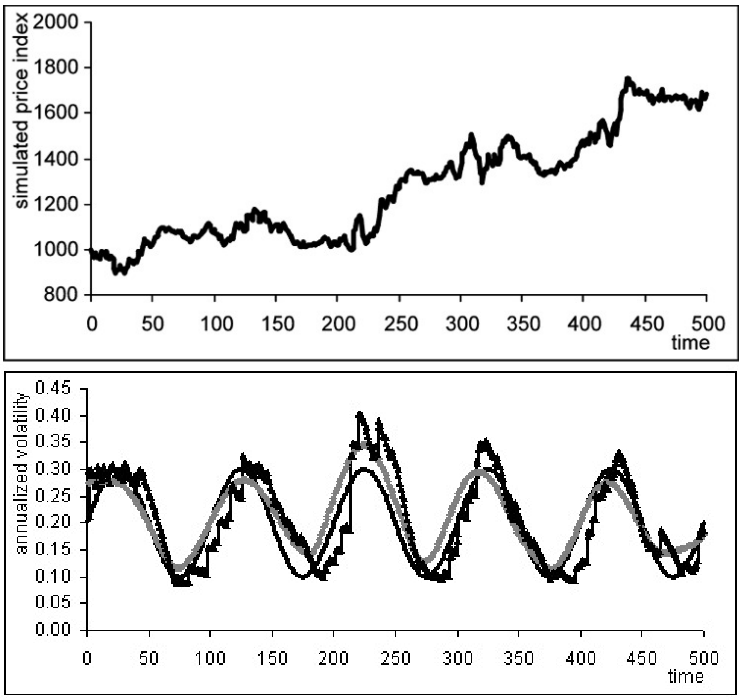


Figure 14: Simulation path for price process $P_t = P_{t-1}\, e^{X_t}$ combined with regression model (20) (top) and annualized volatility estimates for log-returns X_t: $\hat{\sigma}(t)$ (rhombuses, grey) and $\hat{\sigma}_1(t)$ (triangles, black) of the predefined volatility $\widetilde{\sigma}^a(t)$ (continuous line, black) for 500 design points with CV-optimal bandwidths ($h^{CV} = 34$, $h^{CV_{(1)}} = 29$).

C) We finish our simulation studies with a forecasting experiment, continuing on the price process (49) and its realisation displayed in figure 14, i.e. realized innovations $\{\varepsilon_t\}_t$ were frozen. We imagine, that we have observed only the first 251 prices ($P_0, \ldots, P_{250}$) up to a forecast starting point $t_0 = 250$, and call that half of the sample as 'in-sample' part. The second half, called 'out-of-sample' part, will follow within the next year. We plan to forecast the distribution of the 1-day ahead return X_{t+1} with the information available in $t \geq t_0$, successively. For forecasting purposes we have to work with the one-sided setup of the regression model (20) since only past information is available.

The model is calibrated in-sample and optimal parameters are fixed for the out-of-sample part:[65] After estimating a centered return series $\widetilde{R}_1, \ldots, \widetilde{R}_{t_0}$, the optimal smoothing parameter $h_n^{CV_{(1)}}$ is identified for the purpose of non-parametric volatility estimation. Having calculated the series of standard deviation estimates $\hat{\sigma}_{(1)}(1), \ldots, \hat{\sigma}_{(1)}(t_0)$, we estimate the innovation series $\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_{t_0}$ and fit the asymmetric Pearson VII density $f^{VII}_{m_+,c_+,m_-,c_-}$ to nonnegative and absolutes of negative innovations.[66] That way, we use the distribution of X_{t_0} as a forecast of X_{t_0+1}. Afterwards, we develop with the new returns, their volatility estimates and the fixed innovation parametrization the distributions of $X_{251}, \ldots, X_{500}$ successively.

We obtain an optimal one-sided bandwidth $h_n^{CV_{(1)}} = 36$ days. This differs slightly from the optimum on the full sample because of the impact of random innovations in the quite small series. The one-sided annualized, estimated volatilities of the price process are shown in figure 15. The in-sample estimates are grey depicted, while the later pointwise estimators of the out-of-sample part are printed in black. The optimal Pearson VII parameters for estimated innovations $\{\hat{\varepsilon}_t\}_{t=1,\ldots,250}$ in terms of one-sided volatilities and returns are $m_+ = 5.6299, c_+ = 2.9038$ and $m_- = 7.1976, c_- = 3.2758$. They define the right and the left tail of the asymmetric Pearson VII density $f^{VII}_{m_+,c_+,m_-,c_-}$, that are presented in figure 16 relative to the histogram of innovations.

We observe a pretty good approximation of the (scaled) innovation frequencies by the Pearson VII densities (black line) on both sides. Moreover, the illustration presents a comparison to a standardized Gaussian density (in grey):

[65] Alternatively one could recalibrate the model every day or fixed period out-of-sample, to incorporate new parameter information.

[66] Even though we have assumed a symmetric Pearson VII distribution, the random innovations cause different optimal parameters m_+, c_+ vs. m_-, c_- on both sides following the method of moments.

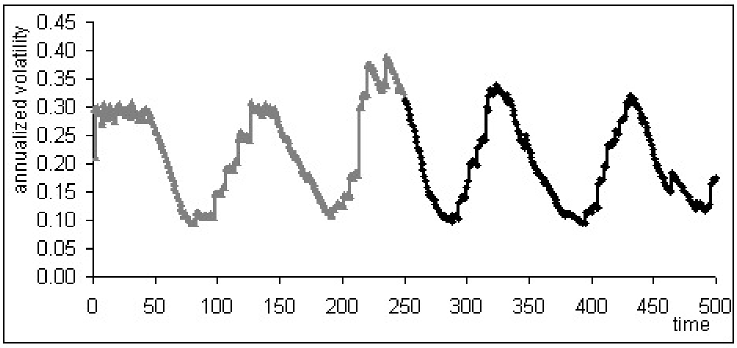


Figure 15: One-sided nonparametric volatility estimates $\hat{\sigma}_1(t)$ in the price process (49) with a CV-optimal bandwidth of 36 days. In-sample estimates are grey, out-of-sample volatilities are black depicted.

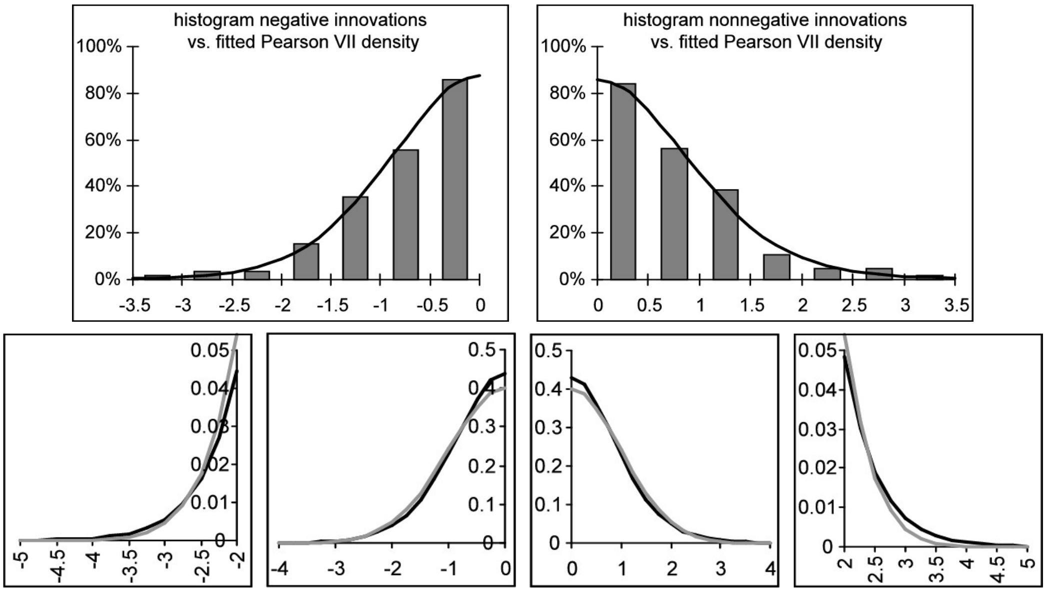


Figure 16: Pearson VII fit of negative and nonnegative innovations by the density $f^{VII}_{5.63,2.90,7.20,3.28}$. The bottom graphs compare the asymmetric Pearson VII (black) and the standard normal density (grey).

Deviations from the normal densities are to be seen in the middle of the distribution and in the tails, where we can conclude that extreme Pearson VII quantiles differ significantly from normal quantiles.[67] There are only slight differences of the left and right distribution fit of innovations due to the simulation setup.

It should be attended, that in general a trade-off between the smoothness of volatility estimates and the innovation's distribution is observed: The bigger the bandwidth in nonparametric volatility estimation the smoother is the volatility graph and the more heavy-tailed are the innovations.[68]

To measure the modelling performance we apply the Kupiec test to shortfall rates of the out-of-sample part. The two-sided hypothesis test is an extension of a binomial test for the likelihood of N shortfalls in a sample of size n, where the true shortfall probability is hypothetical $H_0 : p = \alpha$ for an $(1-\alpha)$ VaR-level. Based on a normal

[67] Concerning the left-sided fit, we see that lower confidence levels (in terms of maximum losses), as 95%, may have Pearson quantiles that are absolutely smaller than the normal (-1.5850 Pearson vs. -1.6449 Gaussian), but for extreme shortfall levels, as 99.5%, Pearson overtakes the Gaussian quantiles (-2.6961 vs. -2.5758).

[68] That is the reason why the Pearson VII fit of innovations may fail for good volatility estimates with small bandwidths. Then the normal distribution may be conservative for the task of extreme quantile approximation.

approximation, Kupiec (1995) developed approximate 95% confidence regions of failure rates. The log-likelihood-ratio

$$LR_p = -2\ln\left[(1-p)^{n-N}p^N\right] + 2\ln\left[\left(1-\frac{N}{n}\right)^{n-N}\left(\frac{N}{n}\right)^N\right] \quad (50)$$

is χ^2_1 distributed under H_0. Thus, the risk measure respectively H_0 is rejected on a 5% level of significance if $LR_p > 3.84$. This test is widely used by financial risk managers to evaluate their risk models, even the penalty zones of the Basel II committee are based on this methodology (see e.g. Jorion (2006)). The Kupiec test is used by scientists as well to support return models, e.g. in Choi and Nam (2008).

Primarily, we use Kupiec backtesting for the non-stationary modelling of the price process (49) to evaluate the forecasted daily returns in terms of certain quantiles. We focus on the loss tail and deduce a relative 1-day Value at Risk $VaR^{1-\alpha,1d}(t)$ for the next day's return X_{t+1} corresponding to the regression model (20):

$$VaR^{1-\alpha,1d}(t) \quad = \quad \bar{X}_t + \hat{\sigma}_{(1)}(t)\, u_{m,c;\alpha^{(1)}} = \bar{X}_t + \hat{\sigma}_{(1)}(t)\, \gamma t_{2m-1;\alpha^{(1)}}\ ,\ \gamma = \frac{c}{\sqrt{2m-1}}, \quad (51)$$

where $u_{m,c;\alpha^{(1)}}$ and $t_{2m-1;\alpha^{(1)}}$ with $\alpha^{(1)} = 2\alpha$ are the quantiles of the left-side Pearson VII fit and the corresponding Student-t expression. This forecast for a maximum loss, that is not exceeded in $(t, t+1]$ with probability $1-\alpha$, is compared to the realized next days return X_{t+1}. A shortfall is observed if $X_{t+1} \leq VaR^{1-\alpha,1d}(t)$. The number of shortfalls in the out-of-sample period is accumulated and evaluated with the Kupiec test statistic LR_p. Figure 17 summarizes for a 99% confidence level of maximum losses. The non-stationary forecasts experienced 5 exceedances over the VaR threshold, which deviates slightly from the expected number of shortfalls, but is within the allowed range of $\{1,\ldots,6\}$ for the 5% level of significance. The model is accepted by the Kupiec test.

LR test (Kupiec-test), 5% level	
VaR confidence level (as 1-α)	99.0%
observations (sample size)	249
number of VaR exceedances	5
shortfall probability p	1.0%
expected exceedances of VaR	2.5
significance level of LR-Test (as 1-β)	95%
c-value χ²-distribution (one-sided)	3.84
test statistic (LR)	1.98
	1.98 < 3.84
Non-stationary model acceptance	yes

Figure 17: Kupiec backtesting of forecasted $VaR^{99\%,1d}(t)$ against realized returns X_{t+1} (out-of-sample) in the non-stationary model.

Furthermore, model acceptance was shown for all possible confidence levels greater-than-or-equal 80%. Most exceedances occured generally, when the volatility was at a low and started again to increase. The example was chosen with quite extreme volatility changes and short periods of the oscillation. Nevertheless the non-stationary approach works. It could be supposed that a parametric VaR based on a long-term standard deviation is advantaged since it covers more periods of the recuring volatility structure. That is why we backtest the basic risk model, introduced in section 2.1, the same way and compare both model performances. Figure 18 gives the number of exceedances for the different approaches. The basic risk model, based on a 250-day standard deviation (fourth column), seems to benefit from the higher average volatility: Less exceedances occur for the most confidence levels and the number of shortfalls is closer to the expectation for levels from 95% to 99.5%. But this VaR-model is too conservative, it is rejected on weaker levels where certain exceedances are required. Based on a window adjusted to the bandwidth of the one-sided NWE (right column) the parametric VaR-model passes the weaker but fails some higher confidence levels.[69]

[69]The basic risk models, based on standard deviations over 50 days and 36 days, respectively, show on the given confidence levels the same number of exceedances.

VaR confidence level (as 1-α)	number of exceedances, model rejection			
	expectation	non-stationary model	normal model (250 days)	normal model (50d, 36d)
80.00%	49.8	45	32	38
90.00%	24.9	29	15	19
95.00%	12.5	16	10	16
98.00%	5.0	6	4	7
99.00%	2.5	5	2	5
99.50%	1.2	3	1	4
99.90%	0.2	1	1	1
99.95%	0.1	0	1	1
99.99%	0.0	0	0	1

Figure 18: Kupiec test results for the non-stationary model and two types of the basic risk model (different horizons of standard deviations) on several confidence levels. Rejections on a 5% level of significance are grey highlighted.

Concluding, the simulation experiment proved that the non-stationary model is not only of theoretical quality. Moreover, it was able to capture simulated return dynamics and to provide satisfactory distributional forecasts as well. Finally, this approach outperformed the basic risk model.

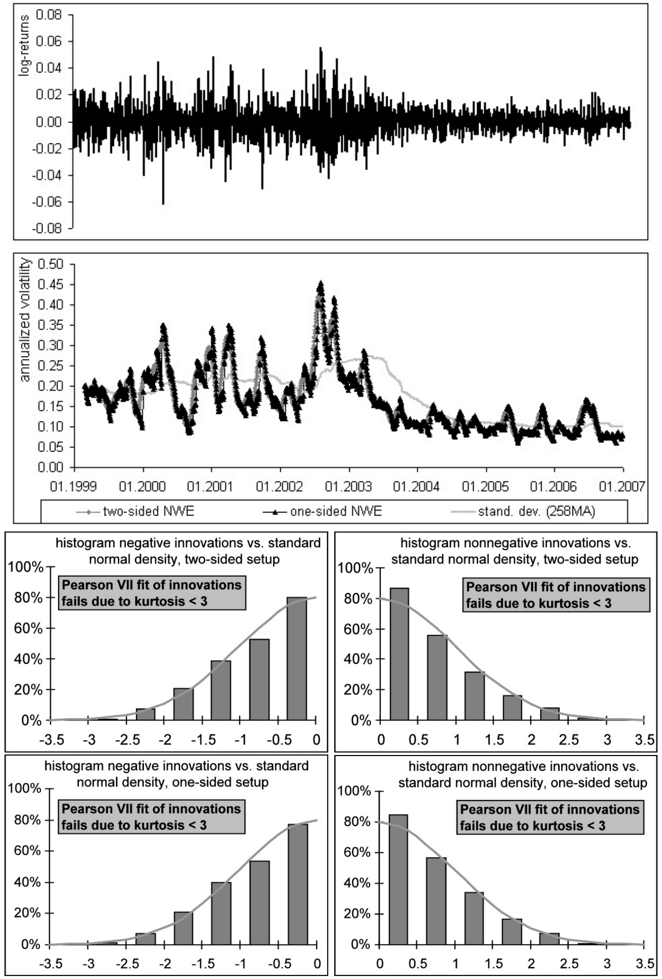
4.2 Empirical studies

In the previous paragraphs we got the confirmation to apply the non-stationary model (20) to real financial time series. The idea of mapping exposures by appropriate benchmarks, introduced within the unfavourable 'basic risk model' in section 2.1, is still promising. We take on the benchmark indices for equity-, interest rate-, credit spread- and currency exposures, that were established in table 1. Returns are determined in the 'preferred conception'. The following studies implement the full non-stationary model univariately to each of the benchmark return series and evaluate the model fit.

A) Our special focus is spent again to the representative examples *EquNA* (MSCI TR Gross North America equity index), *RateUSD* (USD swap rate annual, 5 year), *CredSta* (synthetic BM to model global govt. to swap spread) and *CurrUSD* (ECB Euro Exchange Rate as EUR/USD) of the four exposure types. Their non-stationary modelling is illustrated in figures 19 to 22, following the steps from section 3. After demeaning the log- or diff-returns (top graph), the volatility of the series is estimated nonparametrically. The symmetric, two-sided NWE (21) and the historical, one-sided NWE (22) are presented in the course of time (middle graph).[70] The bottom graphs present the fit of the asymmetric Pearson VII distribution to the residuals, estimated as ratio of demeaned returns and volatility estimators. We distinguish again between the two- and one-sided approach regarding their data input. The tables in each figure report on the corresponding Pearson VII moment estimators and the optimal bandwidths from cross validation.

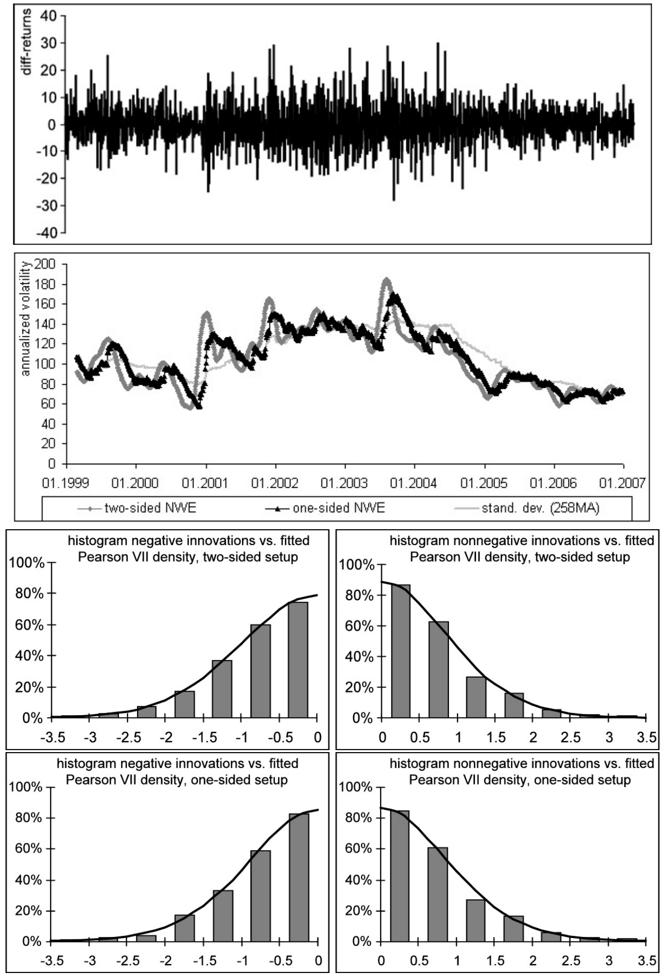
The *EquNA* example reflects the development of one of the major stock markets in the period from 1999 to 2006, with a negative highlight in the burst of the 'I.T. bubble' in the early new millennium. First significant peaks in stock returns appear in March 2000, markets were deflating with full speed in the years 2001 and 2002. The nonparametric volatility estimates immediately react on sequences of extreme log-returns, reaching volatility levels that are more than the double of the long-term average. The one-sided NWE has a certain delay to the both-sided equivalent, but it similarly detects phases of high and low volatility. On the other hand, the empirical standard deviation needs a long until the extreme changes get an impact on the average of 258 centered squared returns. After that financial crisis, the standard deviation declines very slowly while the market volatility was on a low level since summer 2003. The peaked nonparametric volatility graphs are consequence of quite small optimal bandwidths ($h_{opt}^{CV} = 24$, $h_{opt}^{CV_{(1)}} = 30$) from cross validation. The estimators alone catch the market dynamics excellent and the residuals are not as heavy tailed as one may assume. Indeed the Pearson VII fit of innovations fails, since the kurtosis of estimated innovations is lower than 3. In figure 19 alternatively a standard Gaussian density is compared to the histogram of innovations, that sufficiently approximates here and might be adequate in combination with the 'heavy-tailed' volatility estimates.

[70]The volatility series starts at March 1, 1999 to initialize the estimators with two months of past returns. That reduces boundary effects on the left interval end, but does not eliminate them completely if the bandwidth is greater than 40 days. Moreover, boundary corruptions occur on the series' end for the two-sided volatility estimator.



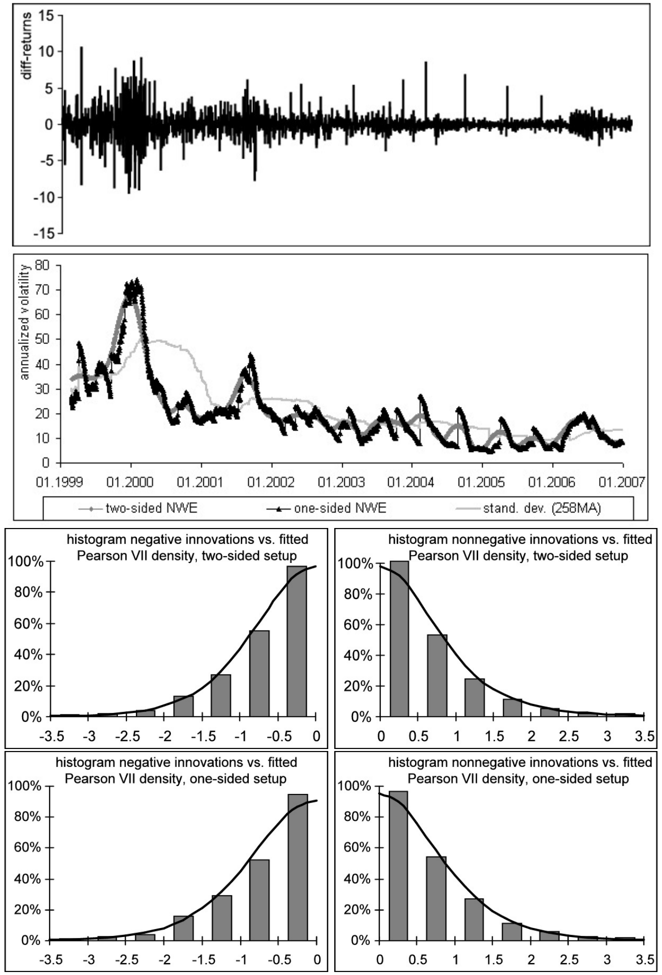
setup	BW h_{opt}^{CV}	m_-	c_-	m_+	c_+
two-sided	24	$n.a.$	$n.a.$	$n.a.$	$n.a.$
one-sided	30	$n.a.$	$n.a.$	$n.a.$	$n.a.$

Figure 19: Top: Daily log-returns of *EquNA* from 1999 to 2006. Middle: Corresponding nonparametric volatility estimators in model (20) compared to standard deviation. Bottom: Left and right Pearson VII fit $f_{m.,c.}^{VII(1)}$ of innovations for the two-sided and one-sided implementation; since the Pearson VII fails for this example, alternatively the Gaussian density is overlaid.



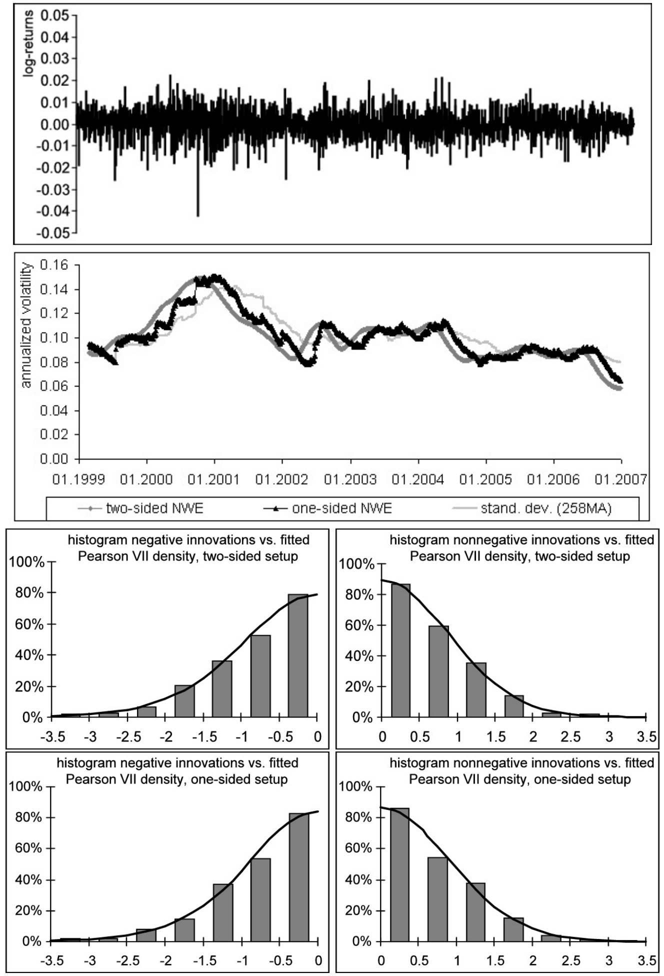
setup	BW h_{opt}^{CV}	m_-	c_-	m_+	c_+
two-sided	41	15.2753	5.4369	7.8298	3.3805
one-sided	120	7.1818	3.3445	6.0330	2.9925

Figure 20: Top: Daily diff-returns of *RateUSD* from 1999 to 2006. Middle: Corresponding nonparametric volatility estimators in model (20) compared to standard deviation. Bottom: Left and right Pearson VII fit $f_{m.,c.}^{VII(1)}$ of innovations for the two-sided and one-sided implementation; attend that the loss tail for a fixed income portfolio is the right side.



setup	BW h_{opt}^{CV}	m_-	c_-	m_+	c_+
two-sided	63	4.7266	2.3271	2.8040	1.6532
one-sided	51	5.8272	2.8017	3.4236	1.9443

Figure 21: Top: Daily diff-returns of *CredSta* from 1999 to 2006. Middle: Corresponding nonparametric volatility estimators in model (20) compared to standard deviation. Bottom: Left and right Pearson VII fit $f_{m.,c.}^{VII(1)}$ of innovations for the two-sided and one-sided implementation; again the loss tail for a credit portfolio is the right side.



setup	BW h_{opt}^{CV}	m_-	c_-	m_+	c_+
two-sided	100	6.2818	3.2537	27.3716	6.5059
one-sided	178	5.5069	2.9202	20.1207	5.7119

Figure 22: Top: Daily log-returns of *CurrUSD* from 1999 to 2006. Middle: Corresponding nonparametric volatility estimators in model (20) compared to standard deviation. Bottom: Left and right Pearson VII fit $f_{m.,c.}^{VII(1)}$ of innovations for the two-sided and one-sided implementation; the loss tail of a currency portfolio is the left.

For the following examples the Pearson VII innovation fit works and we can present the full non-stationary implementation (20). The US 5-year swap rates (figure 20) show three phases, with low volatilities in the first and the last two years. Shocks in interest rates accompany the financial and economic crises from 2001 to 2004, where the 5-year swaps declined from more than 7% to a low of 2.5% p.a. as a consequence of the FED policies to reanimate US economy with low funding rates. Those shocks increased the volatility level from about $60bp$ (based on diff-returns of rates expressed in bp) to about $180bp$ p.a. Here, the optimal one-sided bandwidth was adjusted manually to $h_{opt}^{CV_{(1)}} = 120$, since the CV-function was almost constant for bandwidths greater than 100 with minima in the ranges $[115, 145]$ and $[185, 200]$.[71] The Pearson VII fit of innovations reflects the asymmetry of the return distribution, where the right tail (that is the loss tail for fixed income values) is heavier, as to be seen on the lower Student's df $2m - 1$.

The *CredSta* series (figure 21) fluctuated strongly from summer 1999 to spring 2000, decreasing finally from a swap spread of $-35bp$ to $-65bp$. The volatility estimators start up to a level of $74bp$ annualized volatility, which is more than three times the full sample average. A second maximum is arrived in September 2001 where a flight into the save haven of treasuries after the 9/11 terrorist attacks temporarily slashed the rates. Single peaks in the years 2004 and 2005 cause a sawtooth structure in the one-sided NWE, i.e. a rapid increase and a quite smooth decline (similar to an exponential). The Pearson VII fit of innovations works well, the credit changes turn out to be very heavy tailed, especially in the right tail, expressed by low optimal values of $m.$ and $c.$. The return distribution is leptokurtic and right skewed, i.e. extreme credit losses due to sudden spread expansions are considered with the non-stationary approach.

The last example is the exchange rate EUR to USD in figure 22. For the two-sided volatility estimation the optimal bandwidth has to be manually adjusted ($h_{opt}^{CV} = 100$) with respect to other horizons, other currencies and the one-sided equivalent. The annualized volatilities evolve most time in a range from 8% to 15%, but clusters of higher returns are detected by the NWEs during the years 2000 and 2001. The Pearson VII approximation of the residuals leads to an asymmetric distribution, that is heavier on the left side which is the loss tail of the EUR-investor.

Finally, we extend this analysis to all benchmark return series from 1999 to 2006. Figure 23 reports on the optimal bandwidths for nonparametric volatility estimation and on the Pearson VII parametrization (if existent) of the estimated residuals. The first lines of the table show the two-sided implementation. The one-sided parametrization, based on historical data only, is shown below. As highlighted grey in the tableau, some bandwidth optimizations fail due to a plane (or slowly declining) shape of the CV-function, that inhibits detecting a global minimum. Generally, the bandwidths were restricted to 200 days for reasons of heteroscedasticity and boundary effects, especially for the subsequent smaller samples. Manual bandwidth adjustments had to be conducted to the minor of series. We derived these optima by studying the CV-function, considering similar benchmarks, other time horizons or the opposite (one-/ two-sided) appropriate optima. Consequently, we did not smooth by eyes, but rather decided on quantitative criteria.

Several innovation series in figure 23 are not heavier than a normal distribution and the Pearson VII fit fails since their kurtosis is lower than 3. Due to the separate approximation of negative and nonnegative innovations there are examples where at least one tail can be modelled as Pearson distributed rv. We observe slightly more successes in the one-sided innovation modelling. It is noticeable that the Pearson VII method fails many times, when the volatility estimates are based on small (CV-optimal) bandwidths. We developed the following explanation: The bigger the bandwidth in nonparametric curve estimation the smoother is the volatility estimate and the more heavy-tailed are the innovations. Thus, a perfectly calibrated volatility estimator with a small bandwidth may imply an innovation distribution that is weaker than Gaussian and the Pearson VII fit fails in the non-stationary model. On the other hand, oversmoothed volatility estimates on the same series could produce heavier tailed innovations and the complete framework holds. Hence, there might be a trade-off between the quality in volatility estimation and a successful Pearson VII fit of innovations. For instance, Drees and Starica (2002) choose in their S&P 500 example a bandwidth manually, that is significantly larger than the CV-optimum we derived for that series. In the next step they observed heavy-tailed innovations and fit them asymmetrically by Pearson VII. As a consequence of our NWEs (with a lower CV-error), our Pearson VII fit fails for positive innovations since those are no more heavy tailed.

B) In the next step we compare the model fit to the four standard examples with respect to three different time setups: 1. long-term horizon 1999 to 2006, 2. time horizon 1999 to 2000 and 3. time horizon 2005 to 2006. The periods were chosen to overlap themselves, so that we can compare two models with different information settings in

[71] The decision on the optimal bandwidth was supported by optima on other swap rates (e.g. *RateEUR*) and the analogy to the two-sided bandwidth. Attend to the later remarks.

Parameters of complete implementation non-stationary model

Two-sided setup	**EquEUR**	**EquexEUR**	**EquNA**	**EquAsP**	**EquEM**	**RateEUR**	**RateUSD**	**RateJPY**	**CredSta**	**CredSwa**
=> NWE volatility:										
CV-optimal BW	26	19	24	35	22	34	41	23	63	75
=> innovation fitting:										
m-	*n.a.*	*n.a.*	*n.a.*	9.5314	*n.a.*	26.4848	15.2753	20.1651	4.7266	7.3859
c-	*n.a.*	*n.a.*	*n.a.*	4.2113	*n.a.*	6.6996	5.4369	5.2953	2.3271	3.3564
m+	*n.a.*	*n.a.*	*n.a.*	*n.a.*	15.7506	6.6866	7.8298	*n.a.*	2.8040	5.7025
c+	*n.a.*	*n.a.*	*n.a.*	*n.a.*	4.4192	3.2924	3.3805	*n.a.*	1.6532	2.8941
One-sided setup										
=> NWE volatility:										
CV-optimal BW	24	18	30	38	32	106	120	25	51	117
=> innovation fitting:										
m-	*n.a.*	*n.a.*	*n.a.*	11.0480	65.8141	12.3610	7.1818	*n.a.*	5.8272	7.3385
c-	*n.a.*	*n.a.*	*n.a.*	4.4531	11.7928	4.4445	3.3445	*n.a.*	2.8017	3.3763
m+	*n.a.*	*n.a.*	*n.a.*	*n.a.*	*n.a.*	5.2567	6.0330	53.4319	3.4236	5.4114
c+	*n.a.*	*n.a.*	*n.a.*	*n.a.*	*n.a.*	2.8549	2.9925	10.2646	1.9443	2.8190
Two-sided setup	**CredAAA**	**CredAA**	**CredA**	**CredBBB**	**CredEM**	**CredHY**	**CurrEUR**	**CurrGBP**	**CurrCHF**	**CurrSEK**
=> NWE volatility:										
CV-optimal BW	80	25	28	83	5	7		24	16	19
=> innovation fitting:										
m-	6.3178	*n.a.*	8.4697	3.3385	*n.a.*	*n.a.*		*n.a.*	*n.a.*	*n.a.*
c-	3.1625	*n.a.*	3.6754	1.9523	*n.a.*	*n.a.*		*n.a.*	*n.a.*	*n.a.*
m+	4.0434	8.4898	25.7181	3.7359	*n.a.*	*n.a.*		*n.a.*	256.4915	*n.a.*
c+	2.1762	3.6155	6.8492	2.0024	*n.a.*	*n.a.*		*n.a.*	22.2233	*n.a.*
One-sided setup										
=> NWE volatility:										
CV-optimal BW	120	33	40	141	14	103		101	21	45
=> innovation fitting:										
m-	5.9376	*n.a.*	14.3047	3.3855	*n.a.*	7.1959		5.4396	*n.a.*	*n.a.*
c-	3.0460	*n.a.*	5.1000	1.8276	*n.a.*	3.2015		2.9060	*n.a.*	*n.a.*
m+	3.9626	9.6422	10.6654	3.5661	*n.a.*	4.4542		62.3075	*n.a.*	*n.a.*
c+	2.1004	3.8218	4.1354	2.0555	*n.a.*	2.5230		10.3192	*n.a.*	*n.a.*
Two-sided setup	**CurrDKK**	**CurrNOK**	**CurrUSD**	**CurrCAD**	**CurrJPY**	**CurrAUD**	**CurrNZD**	**CurrSGD**	**CurrHKD**	**CurrEM**
=> NWE volatility:										
CV-optimal BW	30	41	100	97	55	45	55	126	106	17
=> innovation fitting:										
m-	8.5323	7.6567	6.2818	6.4734	8.2892	12.1496	7.8549	21.0223	29.9300	*n.a.*
c-	3.7541	3.7998	3.3527	3.4973	3.6349	5.2782	3.9188	6.6021	8.0164	*n.a.*
m+	4.5259	17.8446	27.3716	22.7077	6.6031	7.1995	7.7123	10.6278	17.7757	*n.a.*
c+	2.3732	5.0011	6.5059	5.7618	3.1892	2.8562	3.0982	4.0050	5.3440	*n.a.*
One-sided setup										
=> NWE volatility:										
CV-optimal BW	52	47	178	88	59	163	167	135	110	17
=> innovation fitting:										
m-	4.9225	7.1637	5.5069	6.2964	11.3313	5.0402	5.2303	16.3051	25.2519	*n.a.*
c-	2.5754	3.4522	2.9202	3.1971	4.3151	2.7502	2.8721	5.4866	7.0335	*n.a.*
m+	5.1287	62.8801	20.1207	*n.a.*	8.6155	8.1105	8.5277	12.6410	15.4323	*n.a.*
c+	2.7431	10.5449	5.7119	*n.a.*	3.7398	3.2905	3.4501	4.5491	5.0632	*n.a.*

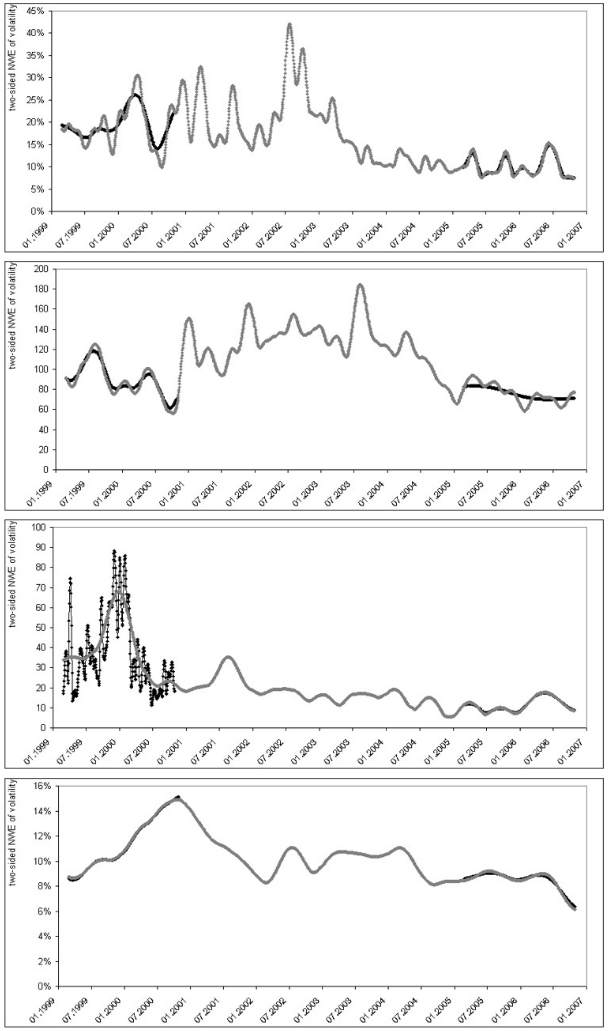
Figure 23: Parameters for the non-stationary modelling of benchmark return series from table 1 for the two-sided and the one-sided implementation. Grey highlighted elements designate manually adjusted bandwidths after CV. Cells with entry '*n.a.*' instead of Pearson VII parameters identify benchmarks where the innovation fitting failed.

the first and last two years. Figures 24 and 25 present the two-sided and one-sided estimated volatility[72] with optimal bandwidths listed in tables below. We observe significant differences in the bandwidth selection during the three periods of time. The smoothing parameters in time-horizons 1999-2006 and 2005-2006 and samples *EquNA*, *CredSta* and *CurrUSD* are still closest to each other. The horizon 1999-2000 shows very different bandwidths concerning most series. Probably, the extremal returns during the financial crises after the y2k had a strong influence on the automatised parameter choice. Consequently, the volatility graphs differ from each other in overlapping periods, especially to be seen in the two-sided implementation for *EquNA* and *CredSta* within the first two years as well as for *RateUSD* on both two-year periods. Strongest deviations in the one-sided estimators appear on *EquNA* and *RateUSD* from 1999 to 2000. All *CurrUSD* volatility estimates are built on quite large optimal bandwidths,[73] so that stronger deviations are not apparently.

Moreover, the optimal bandwidths for all benchmark indices in the three periods of time are listed in figure 26. The above developed result carries forward: The time horizon and extremal returns in one subperiod, respectively, have an important impact on CV-optimal bandwidths and henceforth on nonparametric volatility estimation. Alternatively, time-varying bandwidths, that are reestimated with a certain frequency, could be reconsidered.

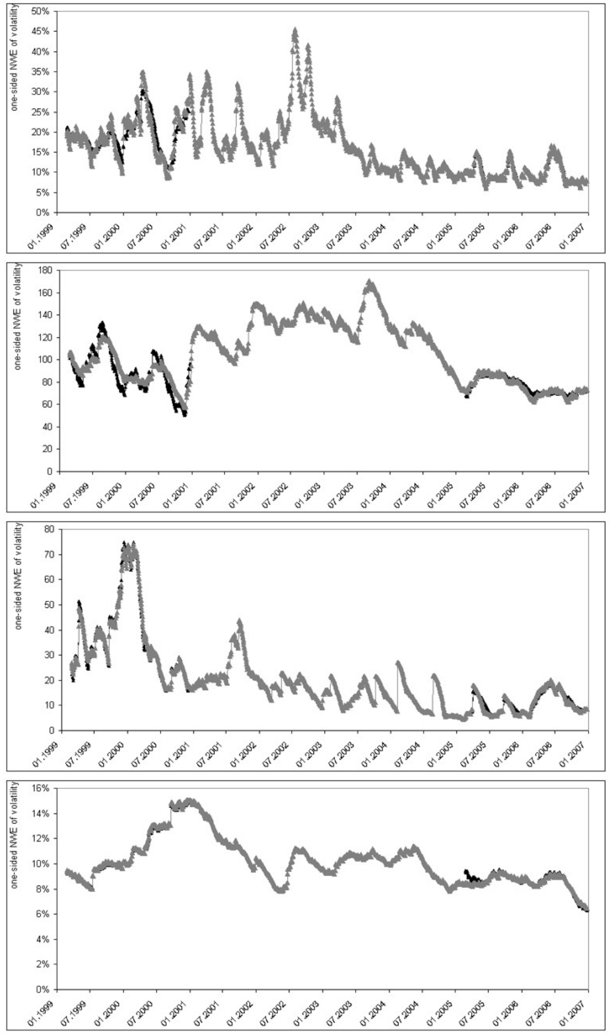
[72]The first and last two months are not estimated in each two-sided setup to reduce boundary effects. Regarding the one-sided implementation only the first two months of data need to be excluded. Some estimation boundary errors remain for bandwidths greater than 40 days.

[73]These are in the both-sided setup twice manually adjusted to 100 days.



BW two-sided setup	horizon 1999-2006	horizon 1999-2000	horizon 2005-2006
EquNA	24	67	34
RateUSD	41	64	200
CredSta	63	8	72
CurrUSD	100	100	118

Figure 24: Nonparametric two-sided volatility estimates in different time horizons: NWEs on the full sample are in grey, NWEs on the two-year subsamples in black. Examples from top to bottom correspond to the order/ parameters in the table below.



BW one-sided setup	horizon 1999-2006	horizon 1999-2000	horizon 2005-2006
EquNA	30	63	34
RateUSD	120	61	192
CredSta	51	44	75
CurrUSD	178	198	159

Figure 25: Nonparametric one-sided volatility estimates in different time horizons: NWEs on the full sample are in grey, NWEs on the two-year subsamples in black. Examples from top to bottom correspond to the order/ parameters in the table below.

Optimal bandwidths for the two-sided and the one-sided setup in different time horizons

setup	EquEUR		EquexEUR		EquNA		EquAsP		EquEM		RateEUR		RateUSD		RateJPY		CredSta		CredSwa	
horizon 1999-2006:	26	24	19	18	24	30	35	38	22	32	34	106	41	120	23	25	63	51	75	117
horizon 1999-2000:	63	45	45	42	67	63	55	50	9	44	116	87	64	61	34	30	8	44	20	34
horizon 2005-2006:	23	24	21	20	34	34	37	36	9	27	200	200	200	192	104	110	72	75	67	70
range	*40*	*21*	*26*	*24*	*43*	*33*	*20*	*14*	*13*	*17*	*166*	*113*	*159*	*131*	*81*	*85*	*64*	*31*	*55*	*83*

setup	CredAAA		CredAA		CredA		CredBBB		CredEM		CredHY		CurrEUR		CurrGBP		CurrCHF		CurrSEK	
horizon 1999-2006:	80	120	25	33	28	40	83	141	5	14	7	103			24	101	16	21	19	45
horizon 1999-2000:	26	26	15	21	31	28	38	37	5	14	10	25			20	41	41	46	32	38
horizon 2005-2006:	62	65	50	52	32	34	92	95	136	140	62	63			54	200	21	199	29	23
range	*54*	*94*	*35*	*31*	*4*	*12*	*54*	*104*	*131*	*126*	*55*	*78*	*0*	*0*	*34*	*159*	*25*	*178*	*13*	*22*

setup	CurrDKK		CurrNok		CurrUSD		CurrCAD		CurrJPY		CurrAUD		CurrNZD		CurrSGD		CurrHKD		CurrEM	
horizon 1999-2006:	30	52	41	47	100	178	97	88	55	59	45	163	55	167	126	135	106	110	17	17
horizon 1999-2000:	36	52	17	35	100	198	80	84	53	54	200	200	200	200	200	200	200	200	22	63
horizon 2005-2006:	47	34	27	40	118	159	179	162	167	158	108	100	116	106	86	198	100	190	10	37
range	*17*	*18*	*24*	*12*	*18*	*39*	*99*	*78*	*114*	*104*	*155*	*100*	*145*	*94*	*114*	*65*	*100*	*90*	*12*	*46*

Figure 26: CV-optimal smoothing parameters on several horizons. In each benchmark column the left (grey) number stands for the two-sided bandwidth while the right is the one-sided. Grey highlighted bandwidths were manually adjusted after CV, furthermore all are restricted to 200 days. Bandwidths differ much between periods of time.

The smoothness of volatility estimates affects the extent of estimated innovations and their distributional fitting. Hence, different horizons bring out different Pearson VII approximations of innovations for the same benchmarks. Figure 27 displays the optimal parametrizations of the four standard examples in the used periods of time. Several times the Pearson VII fit of innovations fails.[74] Again it is to be seen generally, that the smoother the volatility estimate the heavier tailed are the innovations and the smaller are its parameters m_+, c_+, m_-, c_-. Similar effects can be discovered for the other benchmark classes, too.

Pearson VII parameters for the two-sided and the one-sided setup in different time horizons

setup	EquNA								RateUSD							
	m-	c-	m+	c+	m-	c-	m+	c+	m-	c-	m+	c+	m-	c-	m+	c+
horizon 1999-2006:	*n.a.*	*n.a.*	*n.a.*	*n.a.*	*n.a.*	*n.a.*	*n.a.*	*n.a.*	15.2753	5.4369	7.8298	3.3805	7.1818	3.3445	6.0330	2.9925
horizon 1999-2000:	7.2465	3.3699	15.9333	5.2465	7.1698	3.4640	9.8325	3.9547	7.4653	3.5747	6.0688	2.8923	8.0201	3.6565	6.2881	2.9791
horizon 2005-2006:	*n.a.*	*n.a.*	14.7346	5.1682	*n.a.*	*n.a.*	33.3040	7.5373	*n.a.*	*n.a.*	12.6775	4.6470	*n.a.*	*n.a.*	13.8736	4.8844

setup	CredSta								CurrUSD							
	m-	c-	m+	c+	m-	c-	m+	c+	m-	c-	m+	c+	m-	c-	m+	c+
horizon 1999-2006:	4.7266	2.3271	2.8040	1.6532	5.8272	2.8017	3.4236	1.9443	6.2818	3.3527	27.3716	6.5059	5.5069	2.9202	20.1207	5.7119
horizon 1999-2000:	*n.a.*	*n.a.*	*n.a.*	*n.a.*	6.5744	3.1077	5.7159	2.8472	4.0713	2.5145	35.6798	7.2159	4.1676	2.6320	154.142	16.3810
horizon 2005-2006:	6.4910	2.7229	2.7376	1.6334	5.4965	2.6731	3.1447	1.8868	8.5984	4.0097	*n.a.*	*n.a.*	9.5923	4.0262	*n.a.*	*n.a.*

Figure 27: Optimal Pearson VII parameters following the method of moments for standard examples on three time horizons. In each benchmark column the grey coefficients characterize the innovation fit based on the two-sided setup while the black parameters are for the one-sided modelling. Fitting defaults are labeled as '*n.a.*'.

C) We finish this chapter with a forecasting experiment, similar to the simulation setup in section 4.1, part C). Again, we divide the sample series into an 'in-sample' part for model calibration and an 'out-of-sample' part for forecasting and model evaluation by dint of the Kupiec test. We analyse the daily price series of the MSCI North America index from 1999-2002. The first two years are treated as in-sample. On its demeaned log-returns $\widetilde{R}_1, \ldots, \widetilde{R}_{516}$ from 1999 to 2000 the bandwidth $h_{opt}^{CV_{(1)}} = 63$ days was found to be optimally for the nonparametric volatility estimation, compare figures 25 and 26. After having calculated the one-sided NWEs $\hat{\sigma}_{(1)}(1), \ldots, \hat{\sigma}_{(1)}(516)$ and the return residuals $\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_{516}$, the asymmetric Pearson VII density $f^{VII}_{m_+,c_+,m_-,c_-}$ with $m_+ = 9.8325$, $c_+ = 3.9547$, $m_- = 7.1698$, $c_- = 3.4640$ fits the random innovations optimally, compare figure 27. In this vein we implement the regression approach (20) and model the whole return distribution of X_{256}, that is our best prediction for the distribution of X_{257}. According to the 1-day forecasting, we develop the return series and their volatility estimates in the out-of-sample part gradually. Figure 28 presents the in-sample calibration in terms of log-returns, left-side estimated volatilities and the Pearson VII fit of innovations (grey), and already extends the return and volatility series out-of-sample (black).

[74]The Pearson VII method works for more than 60% of the 29 benchmark examples regarding all cases of horizons, two-/ one-sided implementations and separate tail approximations. Due to the number of setups and optimized parameters we do not list all index parametrizations.

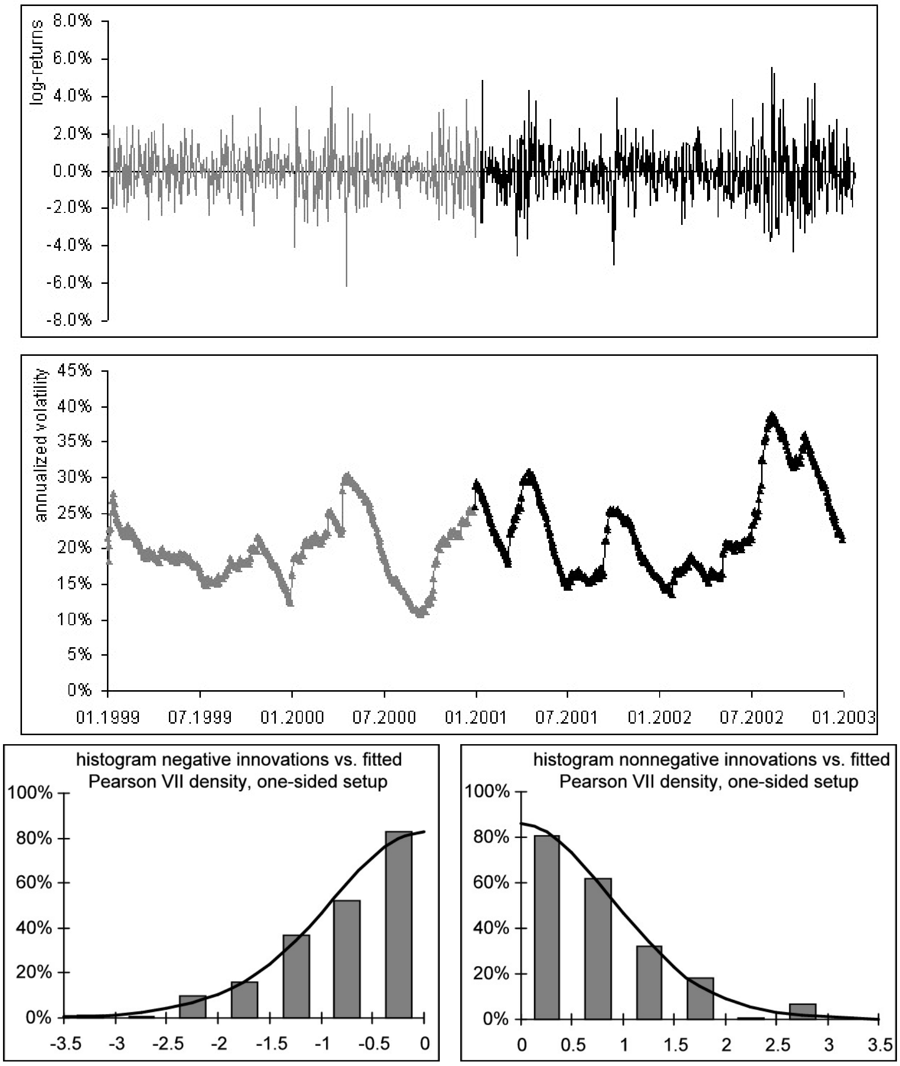


Figure 28: Non-stationary model calibration for the *EquNA* series from 1999 to 2000 (grey, 'in-sample') and model application to the years 2001 and 2002 (black). Top: daily log-returns; middle: nonparametric one-sided volatility estimation with $h_{opt}^{CV_{(1)}} = 63$; bottom: left and right Pearson VII fit of innovations with density $f^{VII}_{9.83,3.95,7.17,3.46}$.

Although the optimal *EquNA* bandwidth differs in the short-term period from the longer optimum, the one-sided NWE is able to capture the stock market shocks during the years 2001 and 2002 very well. A sequence of extreme log-returns increases the volatility estimator to a maximum of 39% p.a. in August 2002. The Pearson VII distribution is leptokurtic and skewed to the left, i.e. extreme losses are more probable than extreme gains.

We assess the quality of the regression model (20) with the Kupiec backtesting, as we compare 1-day forecasted return distributions and certain quantiles, respectively, to the realized returns. We focus on the loss tail and deduce in each point of time $t \geq 516$ a relative Value at Risk $VaR^{1-\alpha,1d}(t)$ on a confidence level $(1-\alpha)$ for the next day's return X_{t+1}, applying formula (51). The number of shortfalls, that are points t where $X_{t+1} \leq VaR^{1-\alpha,1d}(t)$, is accumulated and evaluated with the Kupiec test statistic LR_p from (50). Figure 29 summarizes for a 99% confidence level of maximum losses and a 5% level of significance.

Thereafter the non-stationary implementation of the *EquNA* return series works excellent. Regarding the 513 observations and the 1% proposed shortfall probability, 5.1 exceedances were expected by the backtesting. Indeed the VaR forecast was exceeded 5 times and the non-stationary model is accepted by the Kupiec test. Furthermore, acceptance was derived for all possible confidence levels greater-than-or-equal 80% and with shortfall numbers close

LR test (Kupiec-test), 5% level	
VaR confidence level (as 1-α)	99.0%
observations (sample size)	513
number of VaR exceedances	5
shortfall probability p	1.0%
expected exceedances of VaR	5.1
significance level of LR-Test (as 1-β)	95%
c-value χ²-distribution (one-sided)	3.84
test statistic (LR)	0.003
	0.003 < 3.84
Non-stationary model acceptance	yes

Figure 29: Kupiec backtesting of forecasted $VaR^{99\%,1d}(t)$ vs. realized *EquNA* returns X_{t+1} in the out-of-sample period 2001-2002 for the non-stationary approach.

to the expected exceedances. Figure 30 gives an overview of widely used confidence levels and compares the non-stationary model performance to the basic risk model. The latter modelling failed the Kupiec test for certain levels, since to many exceedance were observed. The regression approach (20) clearly outperformes the classical parametric VaR-model based on 258-day standard deviations.

VaR confidence level (as 1-α)	number of exceedances, model rejection		
	expectation	non-stationary model	normal model (258 days)
80.00%	102.6	111	114
90.00%	51.3	61	66
95.00%	25.7	26	33
98.00%	10.3	12	17
98.50%	7.7	8	15
99.00%	5.1	5	8
99.50%	2.6	3	7
99.90%	0.5	2	2
99.95%	0.3	1	2

Figure 30: Kupiec test results for the non-stationary model and the basic risk model on several confidence levels. Rejections on a 5% level of significance are grey highlighted.

We applied that setup to a plenty of benchmark indices and time horizons with qualitatively similar results. For instance, an analogous survey on exchanges rates EUR/USD from January 4, 1999 to March 31, 2008, with the model being calibrated within the first 6 years using the results from parts A) and B), and being applied to the next 15 months (out-of-sample part) day-to-day, gains a successful real-world approximation, that is supported by Kupiec backtesting.[75] Again this non-stationary implementation outperforms the classical parametric VaR standard.

5 Summary and conclusion

In the first part of the article we introduce a simple parametric Value at Risk approach, as it is used by professional risk managers in banks and insurances, to measure the maximum loss (to a certain probability) of financial returns or a whole portfolio. The VaR calculation is based on an exposure conception for assets with appropriate liquid benchmark indices mapped to it (for equity markets, swap rates of economic areas, credit spreads according to rating classes, important exchange rates to EUR). The benchmark returns are assumed to be iid normal random variables. We get granular on that approach by a comprehensive survey of the modelling assumptions, that have to be rejected in the

[75] This empirical study is available upon request.

majority of hypothesis tests, especially on an 8-year daily data base. The normality of the return distribution is denied as well as the serial identity and independence. The latter results query volatility measurement via empirical standard deviations, in general, and abnegate the time transformation of volatility or VaR to other horizons.

As a consequence a non-stationary, heteroscedastic model for financial returns is developed, based on the ideas of Herzel et al. (2005) and Drees and Starica (2002). In a regression framework the volatility is thought to be exogenous to the process of returns, since no explanatory variables are at hand. Instead the evolution of prices contains the complex market conditions. The vectors of returns are assumed to be independent while having a unconditional covariance structure that changes smoothly in time. With the help of nonparametric regression on equidistant design points we estimate the unconditional (co-)variance directly on centered returns (Nadaraya-Watson estimators, NWEs). Further effort is spent to an accurate modelling of the random residuals. An asymmetric version of the Pearson type VII distribution, that enables heavy tails, is fitted to the estimated innovations. We specialize on an univariate description.

In a technical article on the same non-stationary approach, Gürtler et al. (2009) prove consistency and asymptotic normality for the variance estimates we employ. After outlining these results in terms of the symmetric NWE and of the one-sided NWE, which is based only on recent past data and should be used in forecasting, we deduce requirements for kernels and appropriate bandwidths. The biweight kernel is established. Cross-validation is adopted for an automatised bandwidth selection. The task of fitting innovations via Pearson type VII is simplified by providing a method of moments for parameter estimation and by deriving a connection to the Student-t distribution. By dint of the latter presentation of residuals (and with t-quantiles tabularised) a factor-based VaR calculation can be implemented in terms of the regression modell: The univariate $VaR^{1-\alpha}(t)$ of an exposure $w(t)$ can be modelled as the product of $w(t)$ with a nonparametric estimated volatility $\hat{\sigma}_{(\cdot)}(t)$ and the Pearson VII innovation α-quantile of its benchmark return distribution.

That idea is picked up in a simulation setup and for evaluating the quality of empirical studies for daily log-returns of equities, currencies and diff-returns of interest rates and credit spreads. A Kupiec backtesting confirms the goodness of VaR forecasts that are compared to the next day's returns and evaluated by their shortfall rates. In terms of that backtesting method the non-stationary model clearly outperforms the basic risk model (parametric VaR). Beyond, the simulation studies document how the fit of volatility estimates to a predefined function is improved by a more and more refined data base, and that the non-stationary model is able to capture price processes at all. Extensive empirical studies (for 30 benchmark indices) focus on the practical usage of the nonparametric volatility estimation and its interaction to the Pearson VII fit of innovations.

We observed and do not conceal, that the bandwidth selection via cross validation can not be automatised at all: Rarely the method fails due to finding no accurate minimum in the CV-function. But we provide other quantitative criteria for that choice, so that we avoid smoothing by eyes. Moreover, we observe a trade-off between volatility estimation and Pearson VII innovation fitting: The bigger the bandwidth in nonparametric curve estimation the smoother is the volatility estimate and the more heavy-tailed will be the innovations. Thus, a perfectly calibrated (small) bandwidth may cause that the Pearson VII fit of innovations fails, since the condition of a kurtosis greater than 3 is not satisfied.[76] A failed Pearson VII fit could be compensated conservatively by a Gaussian distribution for innovations. Another observation is, that the optimal parameters (bandwidths and Pearson VII coefficients) change through time. Consequently, a time-varying bandwidth, that is reestimated with a certain frequency, could be a task for further research.

Compared to our main references we think to have developed the following novelties: The non-stationary regression model is applied (univariately) to a variety of financial time series. It is probably used for the first time to approximate credit spreads and to model diff-returns. Our statistical background for nonparametric volatility estimation are consistency and asymptotic normality results from Gürtler et al. (2009). For the Pearson VII modelling of random innovations we provide the formula for parameter estimation based on a method of moments. By dint of a Student-t description of the innovation's distribution we derive a factor-based VaR-presentation of the non-stationary approach, that is easy to implement for a daily real-world execution. The modelling performance is successfully evaluated with the Kupiec backtesting, that furthermore proves an outperformance to the classical, parametric VaR approach. We hope to have succeeded in the splits between theoretical research and practical application.

We finish with some remarks on similar studies: Beside the successful Kupiec test of the regression model, we do not explicitly present the check of all singular model assumptions, mainly concerning the innovation modelling. For that important task we refer back to Drees and Starica (2002), who elaborately proved on a twelve-year S&P 500 return series that the estimated innovations are iid random variables (with similar hypothesis tests as we executed in

[76] On the other hand suboptimal (too smooth) volatility estimates would enable a complete model implementation.

the first part to general returns). They proved the goodness of fit of the asymmetric Pearson type VII distribution for innovations, as they applied a normal transformation of the distribution on realized innovations and normality tests. Moreover, they compared a Student's-t GARCH$(1, 1)$ and a GED EGARCH$(1, 1)$ model, fitted to the same data, with the methods above. They concluded, that the non-stationary return model fits the data significantly better than the conventional GARCH-type models. Finally, they provided a forecasting analysis with 1-, 20-, and 40-days ahead (conditional) distributional forecasts (judged by a normal transformation of that distribution with d-day returns) of the three models, where the non-stationary approach revealed a significantly better performance.

Herzel et al. (2005) showed on a tri-variate example, consisting of the exchange rate EUR/USD, the FTSE 100 index and the 10-year US T-bond rate, that their paradigm describes the changes in the dynamic of the three risk factors well and delivers good multivariate distributional forecasts. The methods are similar to those of the previous paragraph. They proved an outperformance of their forecasted distributions against the $RiskMetrics^{TM}$ (JP Morgan) approach, based on exponential weighted moving average volatilities in a Gaussian return model. They identify the careful modelling of the extremal behavior of innovations as one factor of success, making their approach 'amenable for precise VaR calculations' (Herzel et al. (2005), chapter 8). This is one of the targets, we had on our own agenda. Last but not least, Mikosch and Starica (2004a) extend the non-stationary, nonparametric framework by including a time-varying expected return in the univariate case. They develop the statistical theory for estimating the expected return and volatility consistently. On a S&P 500 daily return series over more than 50 years they give statistical evidence that the expected return as well as the market price of risk varies significantly through time.

Coming back to the widely used VaR-approach of $RiskMetrics^{TM}$, Drees and Starica (2002) as well as Herzel et al. (2005) report on its relationship to the (univariate) non-stationary regression approach. That commercial model, based on a similar multiplicative setup (with return expectation 0 and normal innovations), works factual with a volatility estimate that can be interpreted as Nadaraya-Watson estimator with an exponential kernel.[77] So, this is a special case of our volatility estimation (without demeaning returns, but violating some regularity conditions). Because of our more flexible bandwidth selection and innovation approximation, the non-stationary model is preferred.

Having seen the advantages of the non-stationary model, we can imagine the following fields for future research: The main task will be to develop an adequate multivariate setup for a broad exposure conception. Instead of a direct multivariate setup, a risk aggregation in a simulation approach could be fruitful, e.g. with Cholesky decomposition of correlations, that continues on our univariate factor-based implementation for VaR-purposes. A full nonparametric setup is conceivable, where the (still restrictive) parametric approach for the innovation's distribution could be substituted with a nonparametric kernel density. Following the ideas of Mikosch and Starica (2004a), the inclusion of a time-dependent expected yield, modelled with kernel regression, may be a further step. Based on the basic belief that both recent past and future returns are manifestations of the same unspecified, exogenous economic factors, that evolve smoothly through time, we may use that frame for portfolio optimization in terms of a tactical asset allocation.

References

BLACK, FISHER AND SCHOLES, MYRON, 1973. The pricing of options and corporate liabilities. Journal of political economy 81:637-654.

BOLLERSLEV, TIM, 1986. Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics 31:307-327.

BOLLERSLEV, TIM, ENGLE, ROBERT F. AND NELSON, DANIEL B., 1994. ARCH models. In: ENGLE, ROBERT F. AND MCFADDEN, DANIEL L. (Ed.) Handbook of Econometrics. Elsevier B.V., Amsterdam.

BOUGEROL, PHILIPPE AND PICARD, NICO, 1992. Stationarity of GARCH processes and of some non-negative time series. Journal of Econometrics 52:115-127.

CEIOPS, 2008. Technical Specifications QIS4. Available at (19.05.2008) http://www.ceiops.org/content/view/118/124/Technical Specifications QIS4.doc.

[77]Implementation: $K(x) = a^x I_{[-1,1]}(x)$, $a = \lambda^m$ and $K_h(t) = h^{-1} K\left(\frac{t}{h}\right)$ with $h = m$, following the notation of section 3.1. By default $\lambda = 0.94$ and $m = 74$ for daily returns.

CHAN, NGAI H., 2002. Time Series: applications to finance. Wiley series in probability and statistics. John Wiley & Sons, New York.

CHOI, PILSUN AND NAM, KISEOK, 2008. Asymmetric and leptokurtic distribution for heteroscedastic asset returns: the S_U-normal distribution. Journal of Empirical Finance 15:41-63.

DE HAAN, LAURENS, JANSEN, DENNIS W., KOEDIJK, KEES G. AND DE VRIES, CASPER G., 1994. Safety first portfolio selection, extreme value theory and long run asset risks. In: GALAMBOS, JANOS, LECHNER, JAMES AND SIMIU, EMIL (Ed.). Extreme value theory and its applications. Kluwer Academic Publishers, Dordrecht.

DIEBOLD, FRANCIS X. AND INOUE, ATSUSHI, 2001. Long memory and regime switching. Journal of Econometrics 105:131-159.

DING, ZHUANXIN, ENGLE, ROBERT F. AND GRANGER, CLIVE W.J., 1993. A long memory property of stock market returns and a new model. Journal of Empirical Finance 1:83-106.

DREES, HOLGER AND STARICA, CATALIN, 2002. A simple non-stationary model for stock returns (Preprint). Available at (27.09.2004) http://www.math.uni-sb.de/PREPRINTS/preprint69.pdf.

ENGLE, ROBERT F., 1982. Autoregressive conditional heteroscedastic models with estimates of the variance of United Kingdom inflation. Econometrica 50:987-1007.

ENGLE, ROBERT F., 2004. Risk and Volatility: Econometric Models and Financial Practice. The American Economic Review 94:405-420.

FAMA, EUGENE, 1965. The behavior of stock market prices. Journal of Business 38:34-105.

FAN, JIANQING AND YAO, QIWEI, 2003. Nonlinear Time Series: Nonparametric and Parametric Methods. Springer Series in Statistics. Springer-Verlag, New York.

GASSER, THEO, KNEIP, ALOIS UND KÖHLER, WALTER, 1991. A flexible and fast method for automatic smoothing. Journal of the American Statistical Association 86:643-652.

GRANGER, CLIVE W. J. AND STARICA, CATALIN, 2005. Non-stationarities in stock returns. Review of Economics and Statistics 87:503-522.

GÜRTLER, MARC, KREISS, JENS-PETER AND RAUH, RONALD, 2009. A Non-stationary Approach for Financial Returns with Nonparametric Heteroscedasticity. IF Working Paper Series: IF31, TU Braunschweig.

HAMILTON, JAMES D. AND SUSMEL, RAUL, 1994. Autoregressive conditional heteroskedasticity and changes in regime. Journal of Econometrics 64:307-333.

HERZEL, STEFANO, STARICA, CATALIN AND TÜTÜNCÜ, REHA, 2005. A non-stationary paradigm for the dynamics of multivariate financial returns. In: BERTAIL, PATRICE AND DOUKHAN, PAUL (Ed.). Dependence in Probability and Statistics (Lecture Notes in Statistics, Vol. 187). Springer, New York 2006.

JOHNSON, NORMAN L. AND KOTZ, SAMUEL, 1970. Continuous univariate distributions - 1. Distributions in statistics (Vol. 2). John Wiley & Sons, New York.

JORION, PHILIPPE, 2006. Value at Risk: The New Benchmark of Managing Financial Risk. McGraw-Hill Professional, 3rd edition, New York.

KITAGAWA, GENSHIRO AND NAGAHARA, YUICHI, 1999. A Non-Gaussian Stochastic Volatility Model. The Journal of Computational Finance 2:33-47.

KUPIEC, PAUL, 1995. Techniques for verifying the accuracy of risk measurement models. Journal of Derivatives 3:73-84.

LAMOUREUX, CHRISTOPHER G. AND LASTRAPES, WILLIAM D., 1990. Persistence in variance, structural changes and the GARCH model. Journal of Business and Economic Statistics 8:225-234.

MANDELBROT, BENOIT, 1963. The variation of certain speculative prices. Journal of Business 36:394-419.

MARKOWITZ, HARRY M., 1952. Portfolio selection. Journal of Finance 7:77-91.

MERTON, ROBERT C., 1973. Theory of rational option pricing. Bell Journal of Economics and Management Science 4:141-183.

MIKOSCH, THOMAS AND STARICA, CATALIN, 2004. Non-stationarities in financial time series, the long range dependence and IGARCH effects. Review of Economics and Statistics 86:378-390.

MIKOSCH, THOMAS AND STARICA, CATALIN, 2004a. Stock market risk-return inference. An unconditional non-parametric approach (Preprint). Available at (05.06.07) http://www.math.ku.dk/ mikosch/Preprint/Catalin/jf020904.pdf.

MÜLLER, HANS-GEORG AND STADTMÜLLER, ULRICH, 1987. Estimation of heteroscedasticity in regression analysis. Annals of Statistics 15:610-625.

NELSON, DANIEL B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59:347-370.

SAMUELSON, PAUL A., 1973. Mathematics of speculative prices. SIAM Review 15:1-42.

STARICA, CATALIN, 2003. Is GARCH(1, 1) as good a model as the Nobel prize accolades would imply? (Preprint). Available at (07.03.2005) http://www.math.chalmers.se/~starica/papersoft.pdf.

STRAUMANN, DANIEL, 2004. Estimation in Conditionally Heteroscedastic Time Series Models. Springer, Berlin.

TAYLOR, STEPHEN J., 1986. Modelling financial time series. John Wiley & Sons, New York.