Market Risk Prediction under Long Memory: When VaR is Higher than Expected*

Harald Kinateder

Niklas Wagner

First Version: March 13, 2009

This Version: May 5, 2009

*Harald Kinateder and Niklas Wagner, both at DekaBank Chair in Finance and Financial Control, Passau University. Correspondence: Niklas Wagner, Department of Business and Economics, Passau University, 94032 Passau, Germany. Phone: +49-851-509-3240. E-mail: nwagner@uni-passau.de, nwagner@alum.calberkeley.org

Market Risk Prediction under Long Memory:

When VaR is Higher than Expected

Abstract

Multi-period value-at-risk (VaR) forecasts are essential in many financial risk management applications. This paper addresses financial risk prediction for equity markets under long range dependence. We present the major properties of long memory, its implications for risk management and a novel approach to multi-period market risk prediction under long memory. Our empirical study of established equity markets covers daily index observations during the period January 1975 to December 2007. We document substantial long range dependence in absolute as well as squared returns, indicating a significant influence of long memory effects on volatility. We account for long memory in multi-period value-at-risk forecasts via a scaling based modification of the GARCH(1,1) forecast. Our results show that (i) traditional value-at-risk forecasting techniques underestimate market risk while (ii) our new approach outperforms traditional techniques with as short as 10 or more trading days.

Keywords: long memory, GARCH, fractional Brownian motion, Hurst exponent, autocorrelation, value-at-risk, backtesting, multi-period

JEL Classification: C22

1 Introduction

Several periods of financial market stress (including the market crash of October 1987, a number of accounting scandals at the beginning of the new millennium and the recent 2007/08 banking crisis) have increased both the regulatory as well as the industry demand for effective risk management. A major problem, which risk managers are typically confronted with is to forecast market volatility. To be able to solve this problem, it is necessary to understand the so called "stylized facts" of assets returns. One of these stylized facts is the scaling property denoted as "long memory" or "long range dependence". Long memory implies, that returns in the past affect volatility. Hence, risk managers should take this stylized fact into account to forecast volatility accurately. According to Basel rules[1] of capital adequacy banks must calculate value-at-risk (VaR) numbers for a minimum holding period of *at least* 10 days with a confidence level of 99%. Despite their high practical relevance most focus has been placed on one-day ahead forecasts.

In this paper we provide new insights into risk prediction under long memory by analyzing the long range properties of equity market data. We show that multi-day VaR forecasts can be enhanced significantly by using long memory features. As a result, we develop a new scaling based GARCH(1,1)-LM model for multi-period risk predictions. To our best knowledge, we are the first to improve long-term GARCH volatility forecasts by scaling with the self-affinity index H and by using long memory weights.[2] Despite their high practical importance for risk management scaling-based long memory models are considered sparsely. A few notable exceptions are Embrechts et al. (2005) and Ghysels et al. (2009) who investigate long-term behavior of traditional scaling methods,[3] yet without long memory.

While GARCH models generate volatility forecasts for the very next period, long-term VaR measures usually require volatility predictions for much longer periods of weeks or even several months. Ederington and Guan (2009) evaluate volatility forecasts over longer horizons for non-LM GARCH models and the absolute restricted least squares (ARLS) model. Their results show, that the ARLS is inferior to GARCH type models when index data is considered. The reason for

[1] If one searches the new Basel Accord of the Basel Committe on Banking Supervision (2004) for an instruction for quantitative standards of market risk management systems, one is referred to an earlier version of the Basel Committe on Banking Supervision (1996). It's not necessary to forecast multi-period VaR at once. As such banks are allowed to scale one-day risk measures up to multi-day risk measures.

[2] A detailed description is provided in Section 4.

[3] The square-root-of-time rule is very popular amongst practitioners. Our contribution is to develop a more sophisticated concept, which uses long memory for efficient long-term risk predictions.

this is, that index data is more persistent as for example FX data. Further, the ARLS assumes that log returns are normally distributed. The summary statistics of our data clearly indicates that this is not the case for index data. In addition, Andersen et al. (1999) try to improve DM-$ return volatility forecasts for horizons of five minutes up to one month by using high frequency data. The authors find, that the RSME of all investigated horizons declinces if one considers intraday high frequency data.[4] Our aim is to provide a solution for evaluating daily equity market data, since this frequence is the most available and most used sample frequence in empirical finance.

Especially, in the case of VaR forecasts being estimated over a longer holding period, the influence of long memory increases. Therefore, our contribution is to evaluate long-term equity market risks with a consistent framework. There are several economic, practical and theoretical reasons for this intention. Firstly, the Basel Committee enforces banks to compute their risk metrics for a risk horizon of at least 10 days and nsurance companies must even use a risk horizon of one year. Longer holding periods are desired. Secondly, risk horizons longer than one day are particularly important for financial actors with long time horizons, like pension funds or insurance firms. Several such investors have to rely on longer horizons for long term strategic global asset allocation. Thirdly, the current financial crises demonstrates that long term management is relevant, as financial institutions are not able to improve their equity basis during a crises. Hence, it is necessary that the equity cushion is adequate in *advance* of a stress period such institutions do not run into liquidity demand during a downturn. On the other hand, banks should not use too conservative VaR figures which lead to unnecessary bound money. Berkowitz and O'Brien (2002) investigate the VaR figures of six American banks. They find that the majority uses too high VaR levels. Thus, we evaluate our results not only in terms of conditional coverage.

We test the performance of our scaling based GARCH-LM model by investigating multi-day 99 percent VaR forecasts of four different stock market indices for various horizons τ, $\tau = 5$ (weekly), $\tau = 10$ (biweekly), $\tau = 20$ (monthly) and $\tau = 60$ (quarterly). Our estimates confirm the importance of long memory for risk prediction. Due to our new approach good long-term forecasts are possible, since the deficiencies of GARCH predictions can be cured. All in all, this new approach combines important risk management features, leading to better conditional coverage in stress periods.

[4] The very best sample frequence is not five minutes, but one hour.

The remainder of this paper is organized as follows. Section 2 gives an overview of long memory and the properties of fractional Brownian motion. In addition, this section provides a technique for measuring long range dependence. In Section 3 we present the multi-day forecasting concept of ordinary GARCH(1,1) and introduce our new scaling based long memory GARCH forecasting technique. Furthermore, the underlying VaR framework is presented. Section 4 explains relevant backtesting procedures. Section 5 contains a detailed empirical study. Besides investigating the long memory features of four international stock market indices, we check the forecasting power of our new model. Finally, Section 6 concludes.

2 Literature Review

2.1 Fractional Brownian Motion

A Brownian motion[5] (BM) is a time-continuous stochastic process $(B_t)_{t\in[0,\infty]}$. The increments of BM dB_t are stochastically independent. Time-discrete increments are definied by $\Delta B_t = B_t - B_s \stackrel{d}{=} \mathcal{N}(0; |t-s|), \forall 0 \leq s < t$.

For $0 < H < 1$ fractional Brownian motion (FBM) is a moving average of dB_t, in which past increments of B_t are weighted by the kernel $(t-s)^{H-1/2}$. H denotes the Hurst exponent, which is in fractal geometry known as "self-affinity parameter". In contrast to ordinary Brownian motion, the spreads of the increments can be nearly infinite. The Wiener Brownian motion is the unusual special case, corresponding to the value $H = 0.5$. The increments $\Delta B_t^{(H)} = B_t^{(H)} - B_{t-\tau}^{(H)}$, $\tau > 0$ are stationary, but the process itself is not stationary. $\Delta B_t^{(H)}$ are normal distributed with mean zero and variance

$$\mathbb{E}\Big[\big(B_t^{(H)} - B_{t-\tau}^{(H)}\big)^2\Big] = \tau^{2H}\sigma^2. \tag{1}$$

In the meantime, there are a lot of methods for constructing FBMs available. Besides the original approach of approximating a fractional stochastic integral (Mandelbrot and Van Ness, 1968), the method of Levinson, the wavelet synthesis and the Cholesky decomposition of the covariance matrix are mentionable alternatives. The construction of the latter results from the following strategy: One can simulate sample-paths of FBM by creating a vector of independent standard Gaussian

[5] BM exhibits a constant Hurst exponent $H = 0.5$. Further, BM is sometimes called "Wiener Brownian motion" as Norbert Wiener was the first, who provided the adequate mathematical theory for BM.

numbers $\mathcal{G} = (n_i)_{T\times 1}$. In addition, one must compute a positive definite[6] matrix

$$\Gamma(s,t) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}) \tag{2}$$

and calculate a symmetric square matrix $(A_{i,j})_{T\times T}$ from $\Gamma(s,t) = AA^T$. Finally, the sample path of FBM results from a matrix product $A \times \mathcal{G}$ (Prigarin et al., 2007). The statistical properties and the appearance of FBM are decisively determined by the value of H:

	$\mathbf{0 < H < 0.5}$	$\mathbf{H = 0.5}$	$\mathbf{0.5 < H < 1}$
Autocorrelation	negative (anti-persistent)	zero	positive (persistent)
Dependence between increments	yes	no	yes

Table 1: The impact of H on long memory properties

To understand how the three "standard situations" influence long memory properties, it is useful to analyze various sample paths of FBM. Figure 1 illustrates the sample path of four fractional Brownian motions with different exponents H.

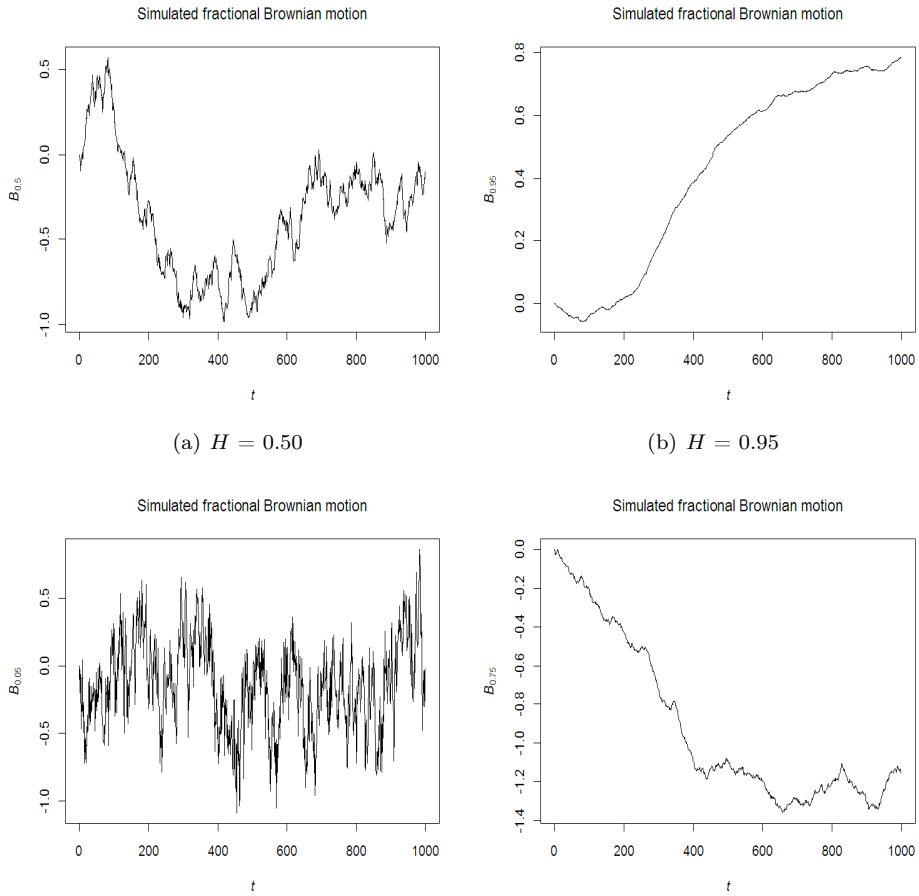
- **$0 < H < 0.5$:** This plot shows no consistent behavior and no trend, at all. Through permanent oscillations there is no trending possible, being typical for ARMA[7] and short memory processes, respectively. All in all, the anti-persistent behavior results of the repeating sign change, because a negative X_t is likely to be followed by a negative X_t.

- **$H = 0.5$:** The Wiener Brownian motion special case of FBM also exhibits no trending, but its sample path is much smoother than the curve for $H = 0.05$. In contrast to $H = 0.05$ there are clearly fewer sign changes eliminating anti-persistence. $H = 0.5$ indicates no persistence.

- **$0.5 < H < 1$:** Both plots ($H = 0.95$ and $H = 0.75$) display high positive autocorrelation leading to highly visible trends. Furthermore, the curves are less fissured, which is characteristic for long memory processes.

[6] It is absolutely necessary, that the input of $\Gamma(s,t)$ generates a positive definite matrix, as without positive definiteness one cannot calculate a square root for $\Gamma(s,t)$.

[7] ARMA processes have no long memory properties and belong therefore to the class of short memory processes.



(a) $H = 0.50$

(b) $H = 0.95$

(c) $H = 0.05$

(d) $H = 0.75$

Figure 1: Sample paths of simulated fractional Brownian motions for $T = 1{,}000$.

Besides graphical techniques, the autocovariance function sheds light on the impact of the Hurst exponent on fractal Brownian motions and financial time series, respectively. Beran (1994) models the autocovariance function (ACVF) of FBM as

$$\gamma_H(\tau) = \frac{\sigma^2}{2}\Big[|\tau - 1|^{2H} - 2|\tau|^{2H} + |\tau + 1|^{2H}\Big]. \tag{3}$$

Additionally, Dieker and Mandjes (2003) demonstrate that for $|\tau| \to \infty$

$$\gamma_H(\tau) \sim \sigma^2 H(2H-1)|\tau|^{2H-2} \tag{4}$$

is valid. Equation (3) and (4) indicate that the autocovariance function of FBM depends on H extraordinary. For $\sigma^2 = 1$ one obtains the autocorrelation function (ACF) $\rho(\tau)$ of fractional Gaussian noise (FGN).[8] Figure 2 demonstrates the impact of H on the ACF of FGN. The plot illustrates amazing results. All ACFs lying

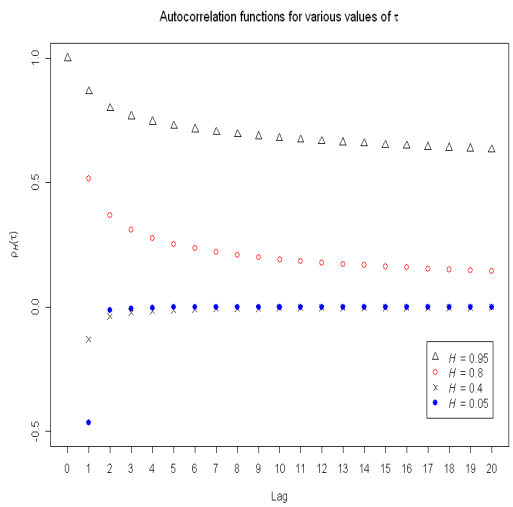
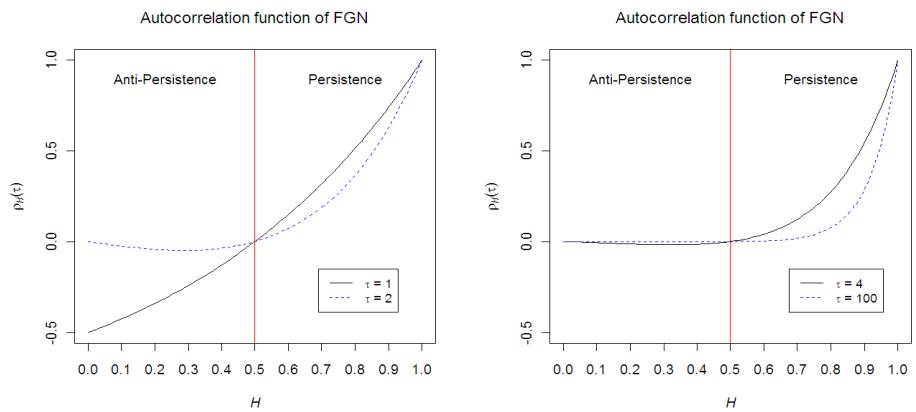


Figure 2: Autocorrelation functions of fractional Gaussian noise for four different Hurst Exponents H

in the long memory area are positive and monotonically decreasing. The more H moves to 0.5, the more $\rho_H(\tau)$ converges to zero. In the special case of $H = 0.5$ no autocorrelation exists, therefore all autocorrelations are zero. If H lies in the anti-persistent region, $\rho_H(\tau)$ converges extremely quick to the abscissa. For instance, $\rho_{0.4}(20) = -0.0022$ and $\rho_{0.05}(20) = -0.0002$. Despite the quick convergence $\rho_H(\tau)$ does not meet the value zero. In contrast to $H \in [0.5, 1]$, all ACFs in the anti-persistent area are lying closely together. These findings can be refined by analyzing $\rho_H(\tau)$ depending on H. Figure 3 provides results for $H \in]0, 1[$ and $\tau \in \{1, 2, 4, 100\}$.

[8] Fractional Gaussian noise refers to the increments $\Delta B_t^{(H)}$ of FBM.



(a) $\tau = 1$ and $\tau = 2$

(b) $\tau = 4$ and $\tau = 100$

Figure 3: Autocorrelation functions of fractional Gaussian noise for four different Lags τ.

The ACF of one lag has the largest spectrum, ranging from -0.5 to 1. Already $\tau = 2$ reduces the range of the spectrum dramatically. Within the anti-persistent area all autocorrelations are negative, but they are very close to zero. This behavior does not change even for very large lags.[9] In the persistent (long memory) region there is a positive relation between H and $\rho_H(\tau)$ identifiable. It's strongest - approximately linear - development is found for $\tau = 1$. Consequently, $\rho_H(\tau)$ converges to zero, if the value of τ rises. To conclude, FBM can capture long memory satisfactory for every time lag τ. So, risk managers can use the concept of FBM to develop sophisticated asset price models. The ACFs of absolute returns indicate informations that should be used in forecasting future volatility. Models which do not use this kind of information are inferior to FBM and related concepts.

2.2 Long Memory

Long Memory or long range dependence, respectively, refers to the slow decay of the autocorrelation function. More precisely, a process contains long memory, if his ACF for $\tau \to \infty$ can be reproduced by a power law

$$\rho(\tau) \sim \frac{c}{\tau^{\delta}}, \tag{5}$$

[9] Between $\tau = 4$ and $\tau = 100$ there is no large difference visible.

with a constant c and a coefficient δ. $\tau \geq 0$ refers to the lag of a financial time series. We focus on continuously compounded returns $(R_t)_{1 \leq t \leq T} = \log(P_t) - \log(P_{t-1})$, based on the financial time series $P_t, t = 0, 1, \ldots, T$. Long memory in finance data is expressed by the slow decay of the ACF of absolute returns ($\pi = 1$) or squared returns ($\pi = 2$): $\rho(\tau) = \text{corr}(|R(t+\tau, \Delta t)|^\pi, |R(t, \Delta t)|^\pi)$. Zumbach (2004) find for absolute returns $\delta \in [0.2, 0.4]$. (5) provides a connection to the Hurst exponent H, as well. Beran (1994), for example, illustrates that equation (5) can be written as

$$\rho(\tau) \sim c\tau^{2H-2}. \tag{6}$$

So, $\delta \in [0.2, 0.4]$ implicates $H \in [0.8, 0.9]$. Generally, $0 < H < 1$ is an indicator for long memory behavior influencing the sign of long range dependence.[10] The impact of H on statistical dependence of time series is displayed in Table 1.

It is well-known that the correlation of empirical asset returns is approximately zero (see e.g. Cont, 2001; Taylor, 2007). Yet, absolute or squared returns behave quite differently. Ding et al. (1993) and Ding and Granger (1996) analyze daily log-returns of Standard & Poor's 500 index from Jan 3, 1928 to Aug 30, 1991. As a major result, they show the slow decay of the ACF in squared and absolute returns.[11] Even for extremely large lags like 3,000 the ACF is significantly different from zero. Also, Engle (1995), Bollerslev and Mikkelsen (1996), and Cont (2001) found highly significant positive autocorrelation in absolute and squared returns of index data. This correlation tells us, that there is information lying in past that can be used to predict future volatility. Yet, Zumbach (2004) remarked that the correlation of asset returns is small, of order 2% to 15%, depending on the time lag and the asset class. As a consequence, it is quite hard to predict future volatility excatly. Thus, risk managers are enforced to use all available information as efficiently as possible. Andersen and Bollerslev (1997) analyse long-run volatility dynamics in high frequency returns. The authors find that long memory is a salient feature of the return generating process, rather than an artifact of regime shifts.

2.3 Long Memory behavior of GARCH(1,1)

Particularly when modeling daily returns, in empirical research the GARCH(1,1) process with conditional normal distributions (Bollerslev, 1986) is the most popular

[10] Additionally, Cont (2005) puts forward economic mechanisms, which may explain long range dependence.

[11] The ACF of absolute returns refers to the autocorrelation of volatility. This slow decay of ACF is sometimes called "hyperbolic". "Exponential" decay refers to short memory processes.

ARCH specification:

$$
\begin{aligned}
R_t &= \mu_t + h_t^{1/2} Z_t && (7)\\
&= \mu_t + \epsilon_t && (8)\\
Z_t &\sim i.i.d.\ \mathcal{N}(0,1) && (9)
\end{aligned}
$$

and

$$h_t = \omega + \alpha\epsilon_{t-1}^2 + \beta h_{t-1} \tag{10}$$

with $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta < 1$. When $\tau \to \infty$, the process σ_t^2 is finite if and only if $\alpha + \beta < 1$, otherwise the process is non-stationary as $\sigma_t^2 \to \infty$. The best way to check if a process contains long memory is to investigate its autocorrelation structure. The ACF of the GARCH(1,1) process for $\tau > 0$ is

$$\rho(\tau) = \rho(1)(\alpha + \beta)^{\tau - 1} \tag{11}$$

with

$$\rho(1) = \frac{\alpha(1 - \alpha\beta - \beta)^2}{(\alpha + \beta)(1 - 2\alpha\beta - \beta)^2}, \tag{12}$$

see e.g. Taylor (2007). Ding and Granger (1996) develop an *approximation* for the theoretical autocorrelation function of GARCH(1,1) models being valid for $\tau \to \infty$:

$$\rho(\tau) \approx (\alpha + \tfrac{1}{3}\beta)(\alpha + \beta)^{\tau - 1}, \qquad \alpha + \beta < 1. \tag{13}$$

The autocorrelation function of short memory processes dies out quickly with exponential decay.[12] Ding and Granger (1996) prove that the ACF of GARCH(1,1) processes belongs to the class of short memory ACFs. This theoretical finding is illustrated in Figure 5. One can clearly observe the fast decay of the ACF of GARCH(1,1) in contrast to the hyperbolic decay of empirical ACFs.

2.4 Non-Scaling based LM-Models

The performance of non-scaling approaches for various horizons is recently investigated by Grané and Veiga (2008). The authors find that long memory models (FIGARCH, HYGARCH) perform better than the short memory GARCH for both error term distributions (Gaussian and Student-t). Yet, backtesting is performed just for 1-day VaR figures. Also Härdle and Mungo (2008) examine long range

[12] Exponential memory is a synonym for short memory. For example, AR, MA and ARMA processes exhibit short memory.

models for short term horizons $\tau \in [1, 5]$.

2.4.1 FIGARCH Model

Baillie et al. (1996) integrate long memory in GARCH volatility by using the filter $(1-L)^d$, with $0 < d < 1$. The combined null $d = 0$ and $d = 1$ indicates no long memory. The so called fractionally integrated GARCH(p, d, q) nests a GARCH $(d = 0)$ with the IGARCH $(d = 1)$ model:

$$\begin{aligned} h_t &= \omega[1-\beta(1)]^{-1} + \left\{1 - [(1-\beta(L)]^{-1}\phi(L)(1-L)^d\right\}\epsilon_t^2 && (14) \\ &= \omega + [1-\beta(1)]^{-1} + \lambda(L)\epsilon_t^2, && (15) \end{aligned}$$

where $\lambda(L) = \lambda_1 L + \lambda_2 L^2 + \cdots$ and $0 \leq d \leq 1$. $0 < d < 1$ reflects long range dependence, since the coefficients in (15) decay hyperbolically. The FIGARCH specification is criticized by Taylor (2007). Taylor remarks, that the variance is infinite for all positive values of d, which is incompatible with the stylized facts for asset returns. Thus, Taylor prefers the FIEGARCH model (Bollerslev and Mikkelsen, 1996, 1999) as it has not this drawback. In addition, Fisher and Calvet (2002) remark that the FIGARCH model is not scale-consistent, because it implies an equivalence between representations of the model at different time scales. The idea of scale-consistency is related to the concept of self-affinity. A random process $\{X(t)\}$ is called self-affine if it satisfies

$$\{X(vt)\} \stackrel{d}{=} \{v^H X(t)\} \tag{16}$$

for all $v > 0$ and $H > 0$. The FIGARCH model provides good one-day volatility forecasts and in general better estimates than direct GARCH. However, if one considers multi-day volatility for long horizons, a problem arises. Due to hyperbolic decay of the coefficients the influence of the coefficients never riches zero, but for long-term forecast horizons the conditional variance tends to converge, which by trend leads to inaccurate volatility predictions.

2.4.2 HYGARCH Model

The hyperbolic GARCH (HYGARCH) is proposed by Davidson (2004). Davidson combines an ordinary GARCH(p, q) with a FIGARCH(p, d, q) model:

$$h_t = \omega[1-\beta(1)]^{-1} + \left\{1 - [(1-\beta(L)]^{-1}\phi(L)[1+\alpha\{(1-L)^d\}]\right\}\epsilon_t^2 \tag{17}$$

Therefore, HYGARCH nests the GARCH ($\alpha = 0$), IGARCH ($\alpha = d = 1$) and FIGARCH model ($\alpha = 1$). Generally, the model is able to reproduce long memory characteristics and volatility clustering, but it suffers from the same drawbacks as the FIGARCH concept.

2.4.3 Other Approaches

Further, Zumbach (2004) accounts for long memory by using a set of historical volatilities. A new family of volatility processes is introduced by measuring the historical volatilities with a set of increasing time horizons. Finally, volatility is forecasted by a simple combination of these historical volatilities. The idea of this concept is to reflect the various horizons of different market participants from intra-day speculators to portfolio managers and pension funds.

2.5 Scaling

In finance scaling is very important, since Basel rules require an underlying time horizon of 10 days for risk predictions. A simply way to get a 10 day risk measure or volatility is to use the square-root-of-time rule. The Basel Market Risk Amendment allows explicitly banks to scale short-term VaR to get an 10-day estimate. Therefore, τ-day VaR is achieved by

$$VaR(1)\sqrt{\tau} = VaR(\tau). \tag{18}$$

The above scaling rule is similar to self-affinity (16) with $H = 0.5$. As we know from FBM, $H = 0.5$ implies an independent time series. It is well known, that financial time series are not independent, because absolute or squared returns are highly correlated. Diebold et al. (1997) investigate the performance of the square-root-of-time for horizons up to one year $\tau = 252$. They truly claim, that the scaling rule relies on 1-day returns being independent and identically distributed (iid) with mean zero and variance σ^2. As a major finding, volatility predictions based on the square-root-of-time rule perform bad, especially for long horizons $\tau = 252$. This finding is not surprising, as advanced scaling methods account for self-affinity properties.[13] Moreover, the authors prefer the Drost and Nijman (1993) formula for temporal GARCH aggregation. In the presence of long memory, this method is not able to include long range dependence in volatility forecasts. As a result, this

[13] Self-affinity suggests not to scale by a fixed exponent $\tau^{H=0.5}$. Instead a fixed exponent one should use the long-memory dependent self-affinity index H.

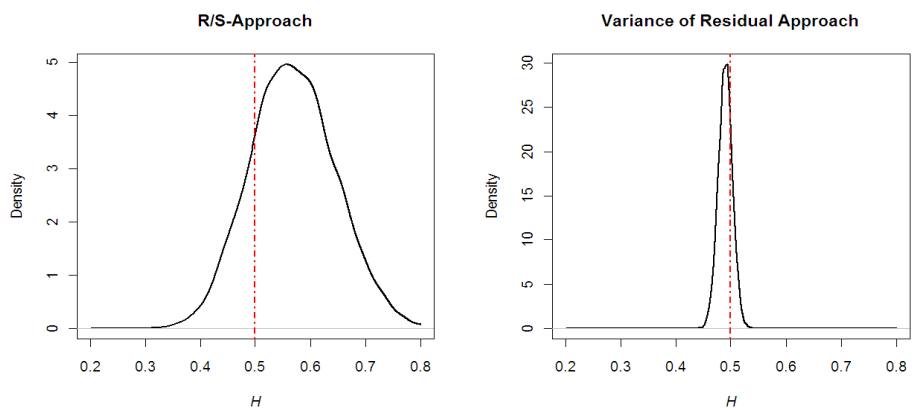
technique is not preferable. Additionally, Daníelsson and Zigrand (2006) detect systematic underestimation of risk, if the underlying assumption of the square-root-of-time rule is incorrect. The degree of underestimation worsens with the time horizon and the confidence level. Embrechts et al. (2005) investigates medium-term risk measures by scaling with $\sqrt{10}$ and provide a framework for one-year ahead risks. The latter is a two step-approach, where in the first step volatility is forecasted by a volatility model up to τ days and then scaled for the rest $T - \tau$ days. If the first forecast is inaccurate, scaling for the rest $T - \tau$ days accumulates the forecasting error increasingly. Caprin (2008) argues untruly that Belratti and Morana (1999) have solved the $\sqrt{T}$ problem. Belratti and Morana just concern short horizons of 1, 5 and 10 days. As we mentioned above, the forecasting error worsens with increasing time horizon. If the data is nearly iid and just a few days are predicted, then the forecasting error is not as bad as in a long horizon environment. Ghysels et al. (2009) compare the the scaling-up method with direct GARCH forecasting (38) and conclude that at short and medium horizons (up to 10 days) they are similar, yet at long horizons the square-root-of-time rule performs better than the direct approach. This finding is quite reasonable, since the conditional variance of GARCH direct forecast converges to its conditional variance when τ rises.[14]

2.6 Estimating Long Memory

In order to measure the sign and the magnitude of (long range) dependence of financial time series we use the Hurst exponent H. In the meantime, the original R/S approach has been extended abundantly to solve the problems associated with this method. Especially, Campbell et al. (1997) point out that the most important shortcoming of the R/S technique is its sensitivity to short-range dependence. Now, there are a lot of alternative methods available. For our study we prefer the concept of detrended fluctuation analysis (DFA), more precisely, the method of Peng et al. (1994) or variance of residuals approach, respectively. This approach is a result of the continuous improvement of original R/S technique, yielding to quite good Monte Carlo outcomes. We demonstrate this progress by estimating the average H of 10,000 independent replications for standard Gaussian noise based on Beran's (1994) estimation technique. So the expected mean is $\overline{H} = 0.5$, since standard Gaussian noise is independent and refers to the increments of Wiener Brownian motion. Ordinary R/S approach is positively biased, since its mean equals 0.569 leading to a bias of 0.069. In contrast to original R/S

[14]For details, compare section 4.2.

technique, Peng's variance of residual approach provides an average estimate of $H = 0.490$, which implies a significantly lower bias. In addition, the root mean square error $RSME = \sqrt{\mathbb{E}[(\widehat{\theta} - \theta)^2]}$ indicates clear dominance of the variance of residuals technique: 0.1034 (R/S approach), 0.0165 (variance of residual approach). Furthermore, the variance of residual concept should be preferred, as its disperse level is considerable lower, see Figure 4. Generally, the Peng technique converts a



(a) $\overline{H} = 0.569$ (b) $\overline{H} = 0.490$

Figure 4: Density of 10,000 estimated Hurst exponent's for independent standard Gaussian numbers. The dashed line refers to the expected mean $H = 0.5$.

bounded random sequence X_t, of length T into m non-overlapping boxes. First of all, one has to compute the cumulative sum X_t:

$$X_t = \sum_{t=1}^{T} \Big(x_t - \frac{1}{T} \sum_{t=1}^{T} x_t \Big). \tag{19}$$

Next, X_t is divided into non-overlapping windows of length m, and a least-squares line $\alpha t + \beta$ is fitted to the cumulated sums. To measure the maximal fluctuations, one considers the square root of an average of squared vertical distances of X_t from the least-squares line:

$$\sigma_t^2 = \left[\frac{1}{m} \sum_{j=1}^{m} \big(X_j - \alpha t - \beta \big)^2 \right]^{1/2}. \tag{20}$$

Self-affine processes can be described by the relation $\sigma_t^2 \propto m^H$, and thus this expression holds for long memory processes, too. Therefore, one can use this ratio

to estimate H by regressioning σ_t^2 on m.

3 Forecasting Equity Market Risk

3.1 Value-at-Risk

Despite demanding no concrete risk management approach, one concept became extremely popular, the value-at-risk (VaR). The VaR is defined as an upper threshold on losses in the sense that these losses are likely to exceed the value-at-risk with only a small probability α. Typically this confidence level is chosen between 1% and 5%. Given some confidence level α and information up to t, $\mathcal{F}_t$, the value-at-risk for the time-interval $(t, t+\tau]$ $VaR^{\alpha}_{t,t+\tau}$ is defined as

$$VaR^{\alpha}_{t,t+\tau} = -\inf\left\{r : F_{t,t+\tau}(r) \geq \alpha\right\}, \tag{21}$$

which yields

$$\mathcal{P}(R_{t,t+\tau} \leq VaR_{t,t+\tau}|\mathcal{F}_t) \geq \alpha.$$

$\mathcal{F}_t$ denotes the information set available up to date t, $F_{t,t+\tau}$ represents the conditional distribution of $R_{t,t+\tau}$. $(R_{t,t+\tau})_{1\leq t\leq T-\tau}=\sum_{i=1}^{\tau}(R_{t+i})$, $\tau \geq 1$, where $R_t\{_{t=1}^{T}\}$ is the one-day log return $R_t = \log(P_t) - \log(P_{t-1})$. This definition implies that the VaR is just calculated for one unit of a special investment.[15] Since the VaR denotes the α-quantile F_{α}^{-1} of the error term distribution F, the VaR simplifies to

$$\widehat{VaR}^{\alpha}_{t,t+\tau} = -(\widehat{\mu}_{t+\tau} + \widehat{h}^{1/2}_{t+\tau}F_{\alpha}^{-1}). \tag{22}$$

Conditional models allow for time-varying conditional mean, which can be considered by additional AR and MA components, that is

$$\mu_t = \mu + \sum_{i=1}^{p} a_i\mu_{t-i} + \sum_{j=1}^{q} b_i\epsilon_{t-j}. \tag{23}$$

If the returns R_t are not autocorrelated, not even on the first lag, and if there is no moving average component, one could neglect conditional mean specifications completely. In this case an unconditional setting $\mu_t = \mu$ is more appropriate. Moreover, equation (22) implies that parametric VaR forecasts need an estimation

[15] As the VaR in (21) only represents one unit of an asset or portfolio, one has to multiply VaR with the total value of the considered item to get its risk position or risk capital, respectively.

of conditional variance h_t and depend on the assumption of the underlying distribution. In the next subsections, we discuss elaborately the GARCH forecasting technique. The distribution of standardized residuals Z_t is often assumed to be standard normal $Z_t \sim \mathcal{N}(0,1)$. Due to heavy tails, for index data this assumption is often violated. In addition, changes in volatility are not symmetric (see e.g. Aboura and Wagner (2009) for a detailed discussion). So, we use an extra distribution, the skewed student-t distribution. For $\nu > 2$ the probability density function $f(Z_t, \xi, \nu)$ of the standardized skewed t-distribution is

$$f(Z_t, \xi, \nu) = \begin{cases} \frac{2}{\xi + \frac{1}{\xi}} g(\xi Z_t | \nu) & if \quad Z_t < 0 \\ \frac{2}{\xi + \frac{1}{\xi}} g\left(\frac{Z_t}{\xi} | \nu\right) & if \quad Z_t \geq 0 \end{cases},$$

where ξ refers to the asymmetry parameter, ν accounts for the tail thickness and $g(.|\nu)$ is a symmetric student density with $\nu > 2$ degrees of freedom

$$g(Z_t|\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu(\nu-2)}} \left(1 - \frac{Z_t^2}{\nu - 2}\right)^{-\frac{\nu+1}{2}}.$$

$\Gamma(\nu) = \int_0^\infty e^{-x} x^{\nu-1} dx$ is the gamma function. The skewed t-distribution nests the symmetric t-distribution for $\xi = 1$ and the Gaussian distribution for $\nu = \infty$ and $\xi = 1$.

3.2 Non-Scaling based GARCH(1,1) forecast

To explain the idea of a τ-day ahead variance estimation of non-scaling based GARCH concepts - whether or not long memory capable - we use a traditional GARCH(1,1) framework. The one-day ahead forecast of conditional variance with respect to equation (9) is given by

$$\mathbb{E}(\epsilon_{t+1}^2 | \mathcal{F}_t) = h_{t+1} = \omega + \alpha \epsilon_t^2 + \beta h_t. \tag{24}$$

$\mathcal{F}_t$ denotes the information set being available up to time t or sub-sigma algebra, respectively. Forecasts being bigger than $\tau > 1$ are more complex and can be

achieved by the law of iterative expectation:

$$\begin{aligned}
\mathbb{E}(\epsilon_{t+\tau}^2|\mathcal{F}_t) &= \mathbb{E}\big[\mathbb{E}(\omega + \alpha\epsilon_{t+\tau-1}^2 + \beta h_{t+\tau-1}|\mathcal{F}_t)\big] && (25)\\
&= \mathbb{E}\big[\omega + \mathbb{E}(\alpha(h_{t+\tau-1}^{1/2} Z_{t+\tau-1})^2 + \beta h_{t+\tau-1}|\mathcal{F}_{t+\tau-2})|\mathcal{F}_t\big] && (26)\\
&= \mathbb{E}\big[\omega + \alpha h_{t+\tau-1} \underbrace{\mathbb{E}(Z_{t+\tau-1}^2|\mathcal{F}_{t+\tau-2})}_{1} + \beta h_{t+\tau-1}|\mathcal{F}_t\big] && (27)\\
&= \mathbb{E}\big[\omega + (\alpha+\beta) h_{t+\tau-1}|\mathcal{F}_t\big]. && (28)
\end{aligned}$$

In the next step, we refine this results by recursive substitution of $h_{t+\tau-1}$ in equation (28):

$$\begin{aligned}
\mathbb{E}(\epsilon_{t+\tau}^2|\mathcal{F}_t) &= \mathbb{E}(\omega + (\alpha+\beta)(\omega + \alpha\epsilon_{t+\tau-2}^2 + \beta h_{t+\tau-2})|\mathcal{F}_t) && (29)\\
&= \mathbb{E}(\omega + \omega(\alpha+\beta) + \alpha\epsilon_{t+\tau-2}^2(\alpha+\beta) + \beta h_{t+\tau-2}(\alpha+\beta)|\mathcal{F}_t) && (30)\\
&= \omega + \omega(\alpha+\beta) + \mathbb{E}(\alpha(h_{t+\tau-2}^{1/2} Z_{t+\tau-2})^2(\alpha+\beta) && (31)\\
&\quad + \beta h_{t+\tau-2}(\alpha+\beta)|\mathcal{F}_t) && (32)\\
&= \omega + \omega(\alpha+\beta) + \alpha h_{t+1}(\alpha+\beta) + \beta h_{t+1}(\alpha+\beta) && (33)\\
&= \omega + \omega(\alpha+\beta) + (\alpha+\beta)^2 h_{t+1}. && (34)
\end{aligned}$$

The mean of equation (28) is $\omega + (\alpha+\beta)h_{t+1}$, so (34) is the same just one period later. Thus, the estimated variance of a $t+\tau$ forecast can be written more compactly:

$$\mathbb{E}(\epsilon_{t+\tau}^2|\mathcal{F}_t) = \omega \sum_{i=0}^{\tau-2} (\alpha+\beta)^i + h_{t+1}(\alpha+\beta)^{\tau-1}, \qquad \tau \geq 2. \tag{35}$$

The first summand of (35) is of special interest, as it can be eased to

$$\omega \sum_{i=0}^{\tau-2} (\alpha+\beta)^i = \underbrace{\frac{\omega}{1-\alpha-\beta}}_{\sigma^2}(1-(\alpha+\beta)^{\tau-1}). \tag{36}$$

σ^2 denotes the unconditional variance of ϵ_t. Inserting (36) in (35) leads to

$$\begin{aligned}
\mathbb{E}(\epsilon_{t+\tau}^2|\mathcal{F}_t) &= \sigma^2(1-(\alpha+\beta)^{\tau-1}) + h_{t+1}(\alpha+\beta)^{\tau-1} && (37)\\
&= \sigma^2 + (h_{t+1} - \sigma^2)(\alpha+\beta)^{\tau-1}. && (38)
\end{aligned}$$

σ^2 denotes the unconditional variance which is equal to $\sigma^2 = \omega/(1-\alpha-\beta)$. If the forecasting horizon τ rises, the conditional variance h_t converges under the restriction $\alpha + \beta < 1$ at an exponential rate fixed by $\alpha + \beta$ to its unconditional variance. As a consequence, all relative weights on past squared returns decline at the same exponential rate whether forecasting volatility for the next day or the distant future Ederington and Guan (2009). These interesting theoretic findings indicates that the GARCH forecasting method is not appropriate, especially for long horizons. Also, Ghysels et al. (2009) criticize direct GARCH forecasting for not being very accurate, if τ is bigger than 10.

3.3 Scaling based GARCH(1,1)-LM: A new Approach

Our approach estimates the τ-day VaR by integrating long memory features. We use the one-day ahead GARCH(1,1) volatility (24) and improve it by the scaled Hurst exponent over τ days and its empirical autocorrelation structure. The reason for this improvement is the strong persistence in index data. Both, the Hurst exponent (Table 3) and the ACF of volatility (Figure 9) indicate long range dependence of daily index data. So, VaR in our setting is given by

$$VaR^{\alpha}_{t,t+\tau} = \mu_{t+\tau} + \phi(t+\tau)F^{-1}_{\alpha}. \tag{39}$$

In contrast to GARCH-based VaR forecasts, we substitute $\sqrt{h_{t+\tau}}$ by a scaling based variable $\phi(t+\tau)$:

$$\phi(t+\tau) = \tau^H \rho_{|R_t|}(\tau)^{H-\rho_{|R_t|}(\tau)}\sqrt{h_{t+1}}. \tag{40}$$

The factor τ^H corresponds to the idea of self-affinity (see equation 16). Further, the square-root-of-time rule implicates $\tau^{1/2}$. The exponent 0.5 can be interpreted as H, indicating Wiener Brownian motion. Our study in section 5 clearly approves that H of absolute index data is significantly different from 0.5. As a result, we integrate this empirical feature by an additional scaling factor τ^H. As a second improvement, we use the slow decay of the empirical autocorrelation structure. $\rho_{|R_t|}(\tau)$ is the autocorrelation coefficient of the time-lag τ. Since $\rho_{|R_t|}(\tau)$ is positive $-$ even for very large lags $-$ this useful information about the variations of volatility can be used for forecasting. There is also a special scaling exponent necessary to ensure that proportion of $\rho_{|R_t|}(\tau)$ is included in the right way. These modifications are enough to get better long horizon volatility predictions.[16] There are no further

[16] Our approach is developed to improve VaR forecasts for horizons being bigger than one day.

improvements of the additionally conditional mean term (23) necessary, since the return process contains no long memory memory, at all.[17]

Due to these modifications, our concept is able to predict VaR figures for considerably long horizons, which can be backtested by comparison with τ-day log returns $R_{t,t+\tau}$. All in all, our model combines traditional volatility concepts with a modified scaling approach. Therefore, we can evaluate risk prospects for various long-term horizons with a consistent framework. In contrast to Christoffersen and Diebold (2000), we find that our new model is able to calculate quite good VaR and volatility figures, respectively, for horizons being considerably bigger than $\tau = 20$.

4 Backtesting VaR

Risk models are only meaningful insofar as they predict risk reasonable well. Therefore, one should test the validity of VaR models through comparison of predicted and actual loss levels. Is the model designed perfectly, the number of observations falling outside the VaR $I = \mathbf{1}_{\{R_{t,t+\tau} < -\widehat{VaR}(t,t+\tau)\}}$ should be in line with the confidence level. All in all, there are two different types of errors: $I > \alpha T$ (type 1 error) and $I < \alpha T$ (type 2 error).[18] The type 1 error describes the situation of rejecting a correct model, while to the type 2 error stands for accepting a misspecified model. To evaluate the out-of-sample performance of the proposed models we follow partly the concept of Kupiec (1995) and Christoffersen (1998). Yet, we do not focus only on likelihood ratio statistics, since in our framework the distance between the actual return and the estimated VaR is meaningful. To check the latter, we use an additional backtesting method.

4.1 Unconditional Coverage

We test for the number of correct exceptions by H_0 : "$\mathbb{E}(I) = \alpha$" versus H_1 : "$\mathbb{E}(I) \neq \alpha$". The test statistic is calculated by

$$LR_{uc} = -2\ln[(1-\alpha)^{T-I}\alpha^I] + 2\ln\{[1-(I/T)]^{T-I}(I/T)^I\}. \tag{41}$$

Under the null hypothesis that α is the correct probability of VaR exceptions, LR_{uc} is asymptotically chi-square distributed with one degree of freedom.

[17] The AR, MA and ARMA processes belong to the class of short memory process, because their ACF dies out quickly. Thus, no further improvements are necessary. In our empirical section this result can be seen, as well. The returns process R_t Hurst exponent is almost 0.5, indicating no long memory (compare Table 3).

[18] α denotes the significance level the VaR is calculated for.

4.2 Test of Independence

In the case of 95 percent VaR, we assume about 13 exceptions each year. These occurrences are expected to be uniformly distributed. If there are 8 violations within two weeks, these increased volatility is not captured by VaR. As a consequence, it is necessary to check, if the VaR exeedances I are statistically independent. The former approach does not take temporal volatility into account, thus an improvement is necessary. Such a test has been developed by Christoffersen (1998). I_{ij} denotes the number of observations in which the state j occurred in one day, while it was at i on the previous observation. If the VaR is exceeded, we set i and j, respectively, to one otherwise to zero. α refers to the unconditional probability of VaR violations, α_i reflects the probability of VaR exeedances conditional on the situation the day before. The corresponding test statistic is

$$LR_{ind} = -2\ln[(1-\alpha)^{I_{00}+I_{10}}\alpha^{I_{01}I_{11}}] + 2\ln[(1-\alpha_0)^{I_{00}}\alpha_0^{I_{01}}(1-\alpha_1)^{I_{10}}\alpha_1^{I_{11}}]. \quad (42)$$

Under the null hypothesis VaR violations are independent and chi-square distributed with one degree of freedom. This means $\alpha = \alpha_0 = \alpha_1 = (I_{01} + I_{10})/T$, which corresponds to the first term in (42). The second term maximizes the likelihood by estimating the conditional probabilities α_0 and α_1.

4.3 Conditional Coverage

The LR_{ind} statistic (42) tests for independence, but it does not take coverage into account. In order to test for correct coverage, Christoffersen (1998) combines both concepts. The combined test statistic

$$LR_{cc} = LR_{uc} + LR_{ind} \quad (43)$$

is the sum of both individual statistics and therefore chi-square distributed with two degrees of freedom.

4.4 Loss Function Based Backtest

The above backtests are suitable for testing, whether the amount of exceedances is in line with the desired coverage rate α and if threshold exceedances occur accidental. In the case of one-day VaR measures these techniques seem to be enough, but for τ-day VaR forecasts one should bare in mind that returns over longer horizons could cause larger losses. As a result, a good VaR forecasting method must cap-

ture these fluctuations appropriately. The former likelihood ratio statistics could indicate a good model, as the coverage rate and independence assumption are not violated. At the same time, huge differences between $R_{t,t+\tau}$ and $VaR_{t,t+\tau}$ can be observed.[19] For this reason, we carry out an additional backtesting approach to ensure that the proposed VaR measures do not differ from $R_{t,t+\tau}$, excessively. This approach has to weight variations from $R_{t,t+\tau}$ in a specific way. A similar loss function based approach is introduced by Lopez (1999). Lopez just evaluates negative threshold exceedances. In our framework, it is important that the estimated VaR is likewise not too weak and too conservative. Berkowitz and O'Brien (2002) investigate risk metrics of six US banks. The authors find that the bank's VaR figures are too conservative, leading to unnecessary bound capital. Due to this, we consider VaR variations in relation to their unconditional expectation α. Yet, we regard VaR exceedances more harmful, therefore the loss function takes the following form:

$$LF(VaR_{t,t+\tau}, R_{t,t+\tau}) =$$

$$\begin{cases} (R_{t,t+\tau} - VaR_{t,t+\tau})^2 & if \quad R_{t,t+\tau} \leq -VaR_{t,t+\tau} \\ |F_{\alpha}^{-1}(R_{1,1+\tau}, ..., R_{t,t+\tau}) - VaR_{t,t+\tau}| & if \quad R_{t,t+\tau} > -VaR_{t,t+\tau} \end{cases}$$

A superior model combines good conditional coverage with a low sample average of LF,

$$\overline{LF} = \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} LF(VaR_{t,t+\tau}, R_{t,t+\tau}). \tag{44}$$

So, one can easily compare different models. On the other hand, the best model never riches $LF = 0$, which would imply no variations. The reason for this is, that at best one detects a coverage rate of exactly α percent. As a consequence, in the best case there is a minimum number of variation. The dimension of this variations is a priori unknown.

5 Empirical Analysis

5.1 Data

We analyze the long range dependence and the performance of both VaR estimation techniques (direct GARCH and scaling based LM) for four international stock market indices. The data is obtained from Thomson ONE Banker. It com-

[19] These differences could be either positive or negative, leading to too low and too conservative VaR measures, respectively.

prises 8,609 daily closing levels P_t from January 1, 1975 to December 31, 2007. In our study, we use non-overlapping continuously compounded percentage returns $R_{t,t+\tau} = [\log(P_{t+\tau}) - \log(P_t)] \cdot 100$ for different sampling frequencies $\tau \geq 1$. Table 2 presents some relevant summary statistics.

All indices exhibit negative skewness. The kurtosis indicates that the returns distribution is fat-tailed, which is typically for index data. We account for these empirical features by using an extra error distribution, namely the skewed student-t distribution.

Index	Frequency	Mean	Std. Dev.	Skewness	Kurtosis
DAX	1	0.0348	1.2129	-0.48	10.78
	5	0.1724	2.6728	-0.62	7.44
	10	0.3414	3.7221	-0.84	7.65
	20	0.6806	5.3957	-1.09	8.24
	60	1.9986	9.7487	-1.06	6.34
DOW JONES	1	0.0357	1.0001	-2.06	57.99
	5	0.1775	2.2092	-1.10	18.09
	10	0.3535	3.0461	-1.16	15.11
	20	0.7048	4.2611	-0.89	8.95
	60	2.0657	7.0707	-0.69	5.82
NASDAQ	1	0.0441	1.1975	-0.31	13.90
	5	0.2199	2.8386	-1.03	10.92
	10	0.4376	4.0783	-1.04	10.20
	20	0.8703	6.0305	-0.93	7.27
	60	2.5652	11.0606	-0.60	4.79
S & P 500	1	0.0356	0.9809	-0.48	41.67
	5	0.1772	2.1720	-0.93	14.64
	10	0.3530	2.9801	-1.03	13.23
	20	0.7050	4.1688	-0.80	8.15
	60	2.0692	6.9038	-0.67	5.74

Table 2: Summary statistics of stock market returns for various sampling frequencies.

5.2 Long Range Dependence

In order to investigate the dependence structure of empirical returns, we calculate estimates of H for various time series and test the null "H_0: $\hat{H} = 0.5$" (no dependence) against the alternative "H_1: $\hat{H} \neq 0.5$" (dependence). Table 3 provides the estimated Hurst exponents and t-values for the mentioned hypothesis. Additionally, we calculate H of simulated Brownian motion consisting 8,608 increments and perform 10,000 replications. The mean of these 10,000 independent replications and the remainder of the results are also provided in Table 3. The null cannot be rejected for any index, hence, our results provide some evidence for the random walk hypothesis of asset returns. However, the random walk hypothesis implies that all non-linear functions of R_t are independent (Andersen and Bollerslev, 1997; Cont, 2001). The latter indicates that all estimated H's for various powers are expected extremely close to 0.5 and the null should not be rejected. Our findings in Table 3 don not satisfy this condition, as the estimated Hurst exponents of absolute and squared returns are significantly different from $H = 0.5$.

Index	R_t	t-value	R_t^2	t-value	$\lvert R_t \rvert$	t-value
DAX	0.520	1.91*	0.769	26.12	0.823	24.01
DOW JONES	0.476	-1.94*	0.614	8.32	0.769	19.99
NASDAQ	0.535	2.83	0.811	17.54	0.846	16.34
S&P 500	0.478	-1.95*	0.632	9.87	0.780	18.72
BM(8608)	0.490	-0.76*	0.487	-0.44*	0.495	-0.23*

Table 3: Empirical estimates of the Hurst exponent H for daily index data from January 1, 1975 to December 31, 2007. A theoretical estimate for simulated ordinary Brownian motion with 8,608 increments is provided for 10,000 replications. * denotes accepting the null at the 95% confidence level.

Further, risk managers are more interested in volatility than in sample returns, so one should focus on the results of squared and absolute values. Stock indices display high autocorrelations in absolute returns (see Figure 9). Approximately for the first 25 lags, there is positive autocorrelation ranging from approximately 0.18 (DOW JONES) to 0.33 (NASDAQ) visible. All indices are located within the strong persistent region $H_{|R_t|} \in [0.75, 1[$ and are highly significant. In contrast to empirical returns, the increments of simulated Wiener Brownian motion ($H = 0.5$) are independent, because in all tested cases the null cannot be rejected.[20] This result is not surprising as rejecting the null would mean that our estimation technique (variance of residuals approach) is failing.

Our previous results in Table 3 indicate high volatility persistence in index data. These findings suggest that the autocorrelation function of $|R_t|$ rather can be described by an ACF with hyperbolic decay as by an approach with exponential decay. In order to investigate how various ACFs fit for different data, we compute two different types of ACFs for selected indices. Firstly, we estimate the long memory ACF (hyperbolic decay) by the following regression model, which is derived from equation (5)

$$\ln(|\rho(\tau)|) = \ln(c) - \delta \ln(\tau) \tag{45}$$

and apply the estimated coefficients $\widehat{c}$ and $\widehat{\delta}$ to equation (5). Secondly, we compute the ACF of a short memory process (11), e.g. GARCH(1,1), for fitted values of α and β. Finally, we compare both ACFs (hyperbolic and exponential decay) with the empirical ACF of volatility $\rho_{|R_t|}$. In order to demonstrate the difference of both concepts, we choose the DAX and NASDAQ index, since they exhibit the

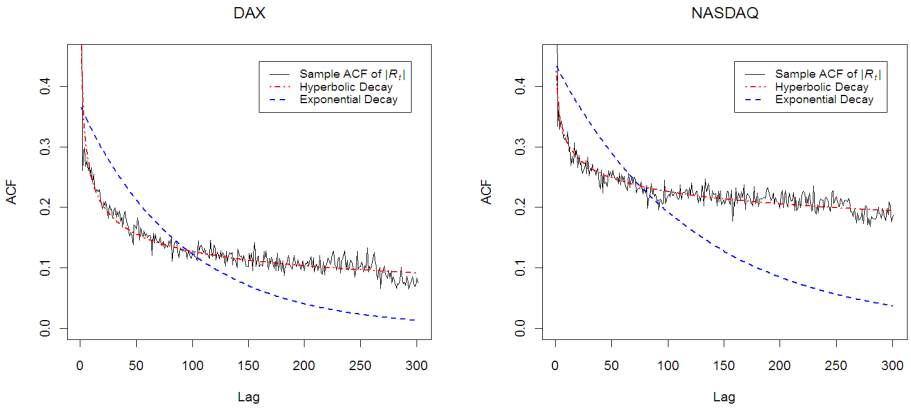
[20] Independence assumes that H for various powers of R_t is 0.5. We document that the null for absolute and squared Gaussian noise cannot be rejected. This stands in sharp contrast to index data (the null is always rejected).

highest autocorrelation in volatility[21] (compare Figure 9). The estimated coefficients are provided in Table 4. The plotted ACFs in Figure 5 confirm our previous

Index	α	β	c	δ
DAX	0.097	0.892	1.344	0.296
NASDAQ	0.099	0.893	1.146	0.137

Table 4: Empirical estimates of the coefficients for ACFs with hyperbolic and exponential decay, respectively.

assumptions. One can observe fast decay of the first 50 lags, followed by persistence[22] up to 300 lags and beyond it. This hyperbolic decay refers to a power law given in equation (5). The ACF with exponential decay does not fit empirical data satisfactory. From $\tau(100)$, the ACF with exponential decay declines to fast, leading to an improper correlation structure.[23] The ACF with hyperbolic decay draws a completely different picture. Thus, both plots approve the previous results in Table 3: there is long memory in daily index data and it can be detected by the Hurst exponent and the slow hyperbolic decay of $\rho_{|R_t|}$. To conclude, both aspects H and $\rho_{|R_t|}$ indicate long range dependence and can be used in the next section to compute scaling based long memory VaR figures for multi-day horizons.



(a) DAX (b) NASDAQ

Figure 5: Sample ACF of absolute returns for the DAX and NASDAQ Composite index.

[21] The autocorrelation of volatility refers to autocorrelation of absolute returns.
[22] Persistence = slow day from $\tau(50)$ till $\tau(300)$.
[23] This explains why the direct GARCH forecast is not suitable for long-term forecasts: the estimated conditional volatility converges with exponential decay time to its unconditional volatility.

5.3 VaR Forecasting Performance Results

In order to examine the out-of-sample performance, we calculate the 1% VaR from January 1, 1991 to December 31, 2007 for different horizons τ, $\tau = 5$ (weekly), $\tau = 10$ (biweekly), $\tau = 20$ (monthly) and $\tau = 60$ (quarterly). All volatility calculations base on a non-overlapping moving window of size one year starting at January 1, 1990 which is updated every τ days. The former data from January 1, 1975 is needed to compute the long range correlation structure as efficiently as possible. In addition, we account for one-lag correlation of R_t by using an AR(1) term in the conditional mean equation. To save space, the AR(1)-GARCH(1,1) and scaling based AR(1)-GARCH(1,1)-LM model just refered as "GARCH" and "GARCH-LM", respectively. The results for both error distributions are provided in Table 5 and 6.

We begin with the results in Table 5. Our estimations confirm the previous assumption that the skewed student-t distribution is clearly the superior distribution when modeling index data. Even for weekly VaR estimates the GARCH-LM method performs somewhat better. With increasing forecasting horizon, the direct GARCH is not able to reach sufficient conditional coverage.[24] By contrast, the GARCH-LM approach shows quite good violation rates for long-term holding periods, as well. Especially for the very important Basel horizon of 10 days, the likelihood ratio statistics indicate good model quality.[25] These findings are exemplarly displayed in Table 6, which shows 10-day ahead VaR results for the DAX und DOW JONES. The remaining plots are available in the appendix.
It is also worth mentioning, that the likelihood ratio statistics of the GARCH-LM approach are not significant at the 1 % confidence level. Just for weekly DOW JONES, biweekly and monthly NASDAQ estimates, the likelihood ratio statistics cannot reject the null at the 5 % confidence level.

The specifications with symmetric normal innovations (Table 6) draw a completely different picture. For almost all horizons, the models with normal innovations underestimate the downside risk considerably.[26] The reason for different VaR ratios between models with normal and skewed t-distribution lies in the differences of the skew parameter ξ. With increasing frequency the skewness behaves not uniformly (see Table 2). For example, the skewness of DOW JONES for sampling

[24] This effect becomes bigger, when τ rises. The reason for this is the fact, that the direct conditional variance GARCH converges to fast to its unconditional variance.

[25] The violation rates rage from 1.13 - 2.25 %.

[26] Almost all LR_{uc} and LR_{cc} statistics indicate that the violation target of 1% is not achieved. However, 10 day ahead LM-VaR of NASDAQ is not significant at the 1% level, but it is significant at the 5% level.

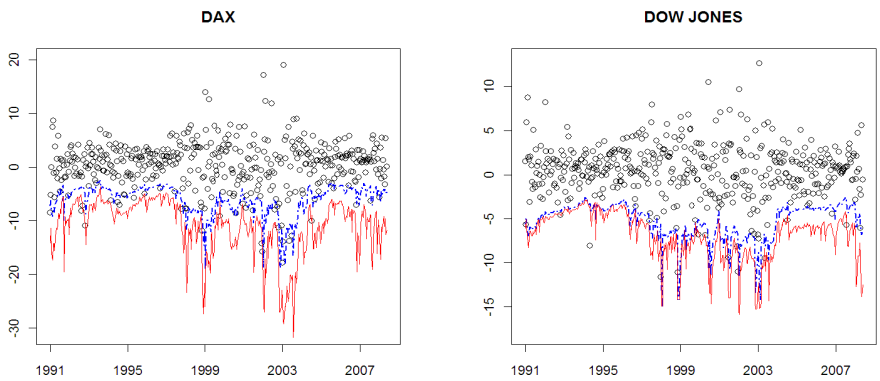


Figure 6: 10-day ahead VaR forecasts for the AR(1)-GARCH(1,1) (blue, dashed line) and AR(1)-GARCH(1,1)-LM (red, solid line) model with skewed student-t innovations from January 1, 1991 to December 31, 2007.

frequency of 10 and 20 days has nearly the same level. However, our findings illustrate that without skewness and fat-tailes there is still an influence of long memory visible, which leads to better multi-period risk predictions than ordinary direct GARCH forecasting. Figure 7 illustrates 10-day ahead forecasts for the DAX and DOW JONES.

Though the LF statistic prefers the GARCH-LM with normal innovation, there is no tough dominance of GARCH in the skewed student-t case. Furthermore, in most cases the levels of LF differ just slightly against the GARCH-LM specification. So, if one compares the LF figures with all aspects, the excellent conditional coverage and graphical fitting, it is not easy to conclude that the GARCH-LM model with skewed student-t innovations is the best of all tested variations (both distribution and forecasting horizon). In addition, the graphical fitting for difficult long horizon VaR predictions seems quite good in comparison to banks internal VaR models.[27] On the other hand, the estimates depend just slightly on the index. The fitting of the NASDAQ is somewhat worse than for the rest of the investigated indices, which can be explained by the higher standard deviation and the higher variations in the correlation structure of the NASDAQ.

All in all, our estimates confirm the importance of long memory for risk prediction. Due to our modified approach, the total number of VaR exeedances in multi-day VaR forecasts can be reduced considerably. This concept provides a rea-

[27]For example, compare Figure 2 in Berkowitz and O'Brien (2002).

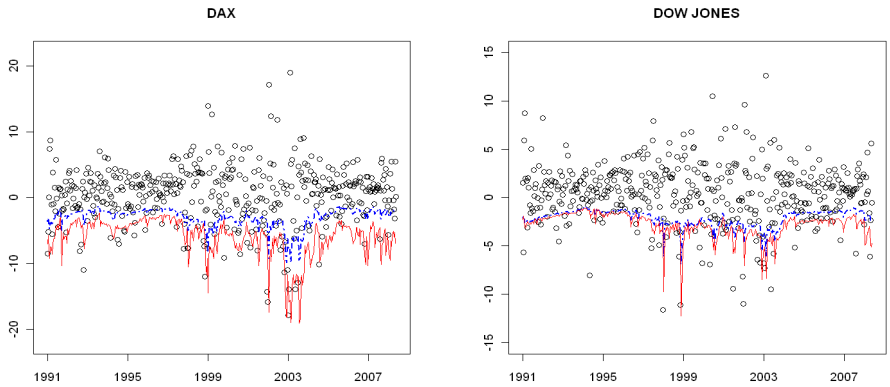


Figure 7: 10-day ahead VaR forecasts for the AR(1)-GARCH(1,1) (blue, dashed line) and AR(1)-GARCH(1,1)-LM (red, solid line) model with normal innovations from January 1, 1991 to December 31, 2007.

sonable solution for the very difficult problem of calculating adequate long-term VaR figures.

5.4 The GARCH(1,1)-LM Approach

In scientific studies non-overlapping τ-day returns are used to ensure that the estimates are independent. Independent returns are necessary to get unbiased backtesting results. This basic issue arises in τ-day risk predictions for $\tau > 1$. Alternatively, one could compute $R_{t,t+\tau}$ every day, which leads to overlapping returns. Overlapping returns cause dependency, which requires special backtesting precautions. Primarily, one has to account for autocorrelation of order $\tau - 1$, if τ-day overlapping returns are used. On the other hand, daily updating reduces huge jumps in overlapping returns, since τ-day non-overlapping returns are more distant when τ rises. Another advantage of daily updating is that backtesting criteria like the Basel traffic light could be achieved easier, because of daily crossovers it is easier to match the VaR with the tails of a daily updated τ-day return distribution.

As a result, practitioners can use this approach by computing VaR figures daily, as well. However, it is important that there is enough data available to get suitable estimates for the long range autocorrelation structure. Moreover, we point out that the estimation technique for H is of vital importance, since it determines the accuracy of the results. Firstly, it is absolutely necessary that the estimated values

of H do not scatter much. The variance of residuals method satisfies this condition excellently (see Figure 4. Secondly, the level of H must be reasonable, too, since too low values of H near by 0.5 neutralize the long memory effect.[28]

6 Conclusion

In this paper we address the problematic of multi-day value-at-risk and volatility forecasts, respectively. One way to solve this difficult issue is to use long memory for computing long term risk figures. Our novel scaling based GARCH-LM concept clearly dominates direct GARCH forecasting.[29] As a consequence, our estimates confirm the importance of long range dependence for VaR prediction. The specification with skewed student-t innovations satisfies the desired Basel confidence level of 99 % extraordinary well, even for quarterly returns.

All in all, the new equity market risk management model provides a consistent framework to evaluate financial risks much longer than $\tau = 10$ days. Therefore, banks using this approach can gain more accurate risk metrics in comparison to those using direct GARCH forecasting or traditional square-root-of-time rule. The latter is not appropriate to forecast volatility, as our estimates clearly indicate that equity market returns are not independent. The square-root-of-time rule explicitly assumes the returns process as independent and identically distributed. Last but not least, the current crises has shown how important long term risk management is. Certainly more future work in this field is needed.

[28] By contrast, too high levels overstate the actual downside risk.

[29] The gap between the original direct GARCH forecast and the GARCH-LM approach becomes bigger, if the forecast horizon rises.

References

Aboura, S. and N. Wagner (2009): Extreme asymmetric volatility, leverage, feedback and asset price, Working Paper, Université de Paris-Dauphine and Passau University.

Andersen, T. G. and T. Bollerslev (1997): Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns, *Journal of Finance* 52: 975–1004.

Andersen, T.G., T. Bollerslev, and S. Lange (1999): Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon, *Journal of Empirical Finance* 6: 457–477.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996): Fractionally integrated generalized autoregressive conditional hetorskedasticity, *Journal of Econometrics* 74: 3–30.

Basel Committe on Banking Supervision (1996): Amendment to the Basel Capital Accord to incorporate market risk, BIS, Basel, Switzerland.

Basel Committe on Banking Supervision (2004): International convergence of capital measurement and capital standards: A revised framework, BIS, Basel, Switzerland.

Belratti, A. and C. Morana (1999): Computing value at risk with high frequency data, *Journal of Empirical Finance* 6: 431–455.

Beran, J. (1994): *Statistics for Long-Memory Processes* (Chapman & Hall/CRC, Boca Raton).

Berkowitz, J. and J. O'Brien (2002): How accurate are value-at-risk models at commercial banks?, *Journal of Finance* 57: 1093–1111.

Bollerslev, T. (1986): Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31: 307–327.

Bollerslev, T. and H. O. Mikkelsen (1996): Modeling and pricing long memory in stock market volatility, *Journal of Econometrics* 73: 151–184.

Bollerslev, T. and H. O. Mikkelsen (1999): Long-term equity anticipation securities and stock market volatility dynamics, *Journal of Econometrics* 92: 75–99.

Campbell, J. Y., A. W. Lo, and A. C. MacKinley (1997): *The Econometrics of Financial Markets* (Princeton University Press, Princeton).

Caprin, M. (2008): Evaluating value-at-risk measures in presence of long memory conditional volatility, *Journal of Risk* 10: 79–110.

Christoffersen, P. (1998): Evaluating interval forecasts, *International Economic Review* 39: 841–862.

Christoffersen, P. F. and F. X. Diebold (2000): How relevant is volatility forecasting for financial risk management, *Review of Economics and Statistics* 82: 12–22.

Cont, R. (2001): Empirical properties of asset returns: stylized facts and statistical issues, *Quantitative Finance* 1: 223–236.

Cont, R. (2005): Long range dependence in financial markets, in Lutton, E. and G. Lévy-Véhel, eds., *Fractals in Engineering* (Springer, Berlin).

Daníelsson, J. and J.-P. Zigrand (2006): On time-scaling of risk and the square-root-of-time-rule, *Journal of Banking and Finance* 30: 2701–2713.

Davidson, J. (2004): Moments and memory properties of linear conditional heteroskedasticity models, and a new model, *Journal of Business and Economic Statistics* 22: 16–19.

Diebold, F.X., A. Hickman, A. Inoue, and T. Schuermann (1997): Converting 1-day volatility into h-day volatility: Scaling by $\sqrt{h}$ is worse than you think, Working Paper, Wharton Financial Institutions Center 97-34.

Dieker, A. B. and M. Mandjes (2003): On spectral simulation of fractional Brownian motion, *Probability in the Engineering and Informational Sciences* 17: 417–434.

Ding, Z., R. F. Engle, and C. W. J. Granger (1993): A long memory property of stock market returns and a new model, *Journal of Empirical Finance* 1: 83–106.

Ding, Z. and C.W.J. Granger (1996): Modeling volatility persistence of speculative returns: A new approach, *Journal of Econometrics* 73: 185–215.

Drost, F. C. and T. E. Nijman (1993): Temporal aggregation of garch processes, *Econometrica* 61: 927–909.

Ederington, L. H. and W. Guan (2009): Longer-term time series volatility forecasts, *Journal of Financial and Quantitative Analysis, Forthcoming* .

Embrechts, P., R. Kaufmann, and P. Patie (2005): Long-term risk management, Working Paper, ETH Zurich.

Engle, R. F. (1995): *ARCH: Selected Readings* (Oxford University Press, Oxford).

Fisher, A. and L. Calvet (2002): Multifractionality in asset returns: Theorie and evidence, *Review of Economics and and Statistics* 84: 381–406.

Ghysels, E., A. Rubia, and R. Valkanov (2009): Multi-period forecasts of volatility: Direct, iterated, and mixed-data approaches, Working Paper, University North Carolina.

Grané, A. and H. Veiga (2008): Accurate minimum capital risk requirements: a comparison of several approaches, *Journal of Banking and Finance* 32: 2482–2492.

Härdle, W. K. and J. Mungo (2008): Value-at-risk and expected shortfall when there is long range dependence, Working Paper, Humboldt-University zu Berlin.

Kupiec, P. (1995): Techniques for verifying the accuracy of risk measurement models, *Journal of Derivatives* 3: 73–84.

Lopez, J.A. (1999): Methods for evaluating value-at-risk models, *Federal Reserve Bank of San Francisco Economic Review* 2: 3–17.

Mandelbrot, B. B. and J. W. Van Ness (1968): Fractional brownian motion, fractional noises and applications, *SIAM Review* 10: 422–437.

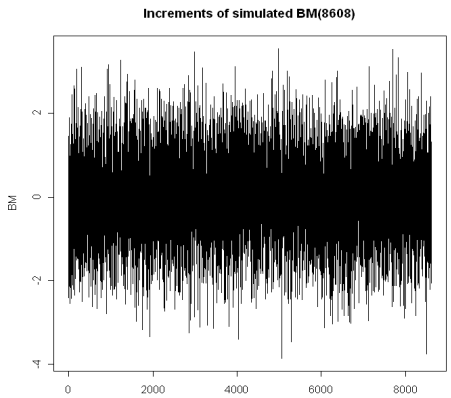
Peng, C.-K., S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger (1994): Mosaic organization of dna nucleotides, *Physical Review E* 49: 1685–1689.

Prigarin, M., S., K. Hahn, and G. Winkler (2007): Evaluation of two numerical methods to measure the hausdorff dimension of the fractal brownian motion, Working Paper, Institute of Biomathematics and Biometry, Munich.

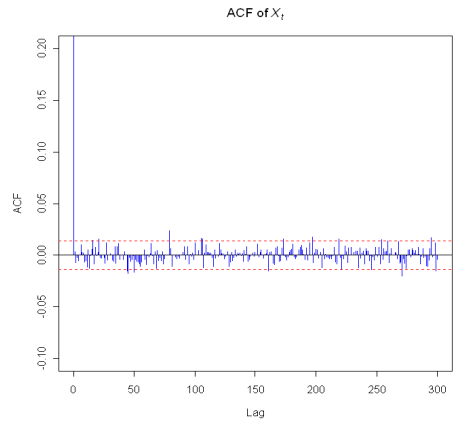
Taylor, S. (2007): *Asset Price Dynamics, Volatility and Prediction* (Princeton University Press, Princeton).

Zumbach, G. (2004): Volatility processes and volatility forecast with long memory, *Quantitative Finance* 4: 70–86.

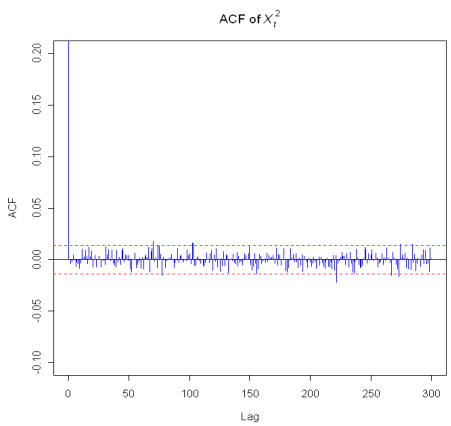
7 Appendix



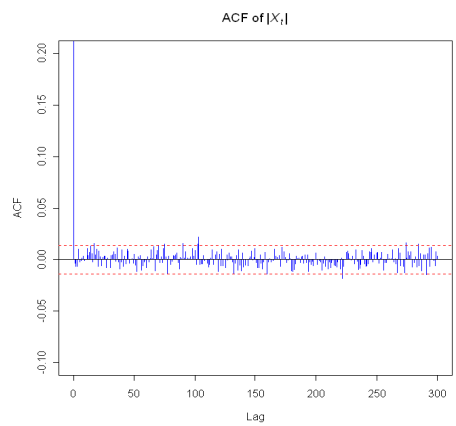
(a) Increments X_t



(b) ACF of increments X_t



(c) ACF of squared increments X_t^2



(d) ACF of absolute increments $|X_t|$

Figure 8: Increments and ACFs of simulated ordinary Brownian motion for various powers.

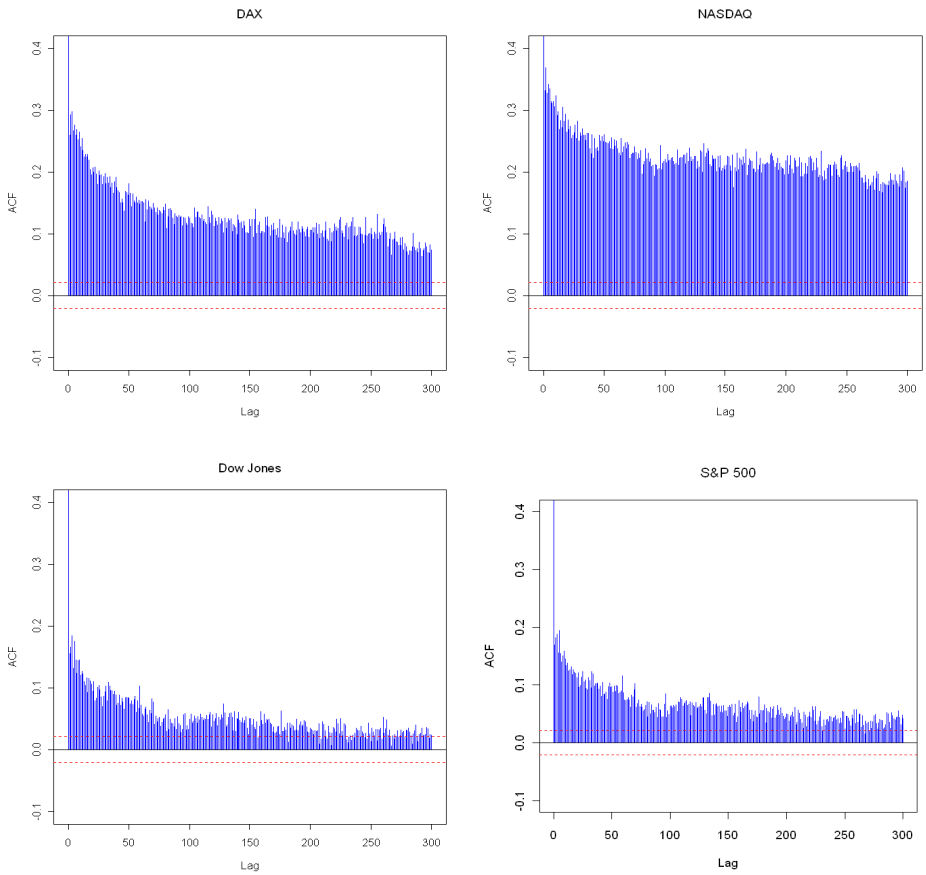


Figure 9: ACFs of absolute index returns. Sample period: January 1, 1975 to December 31, 2007.

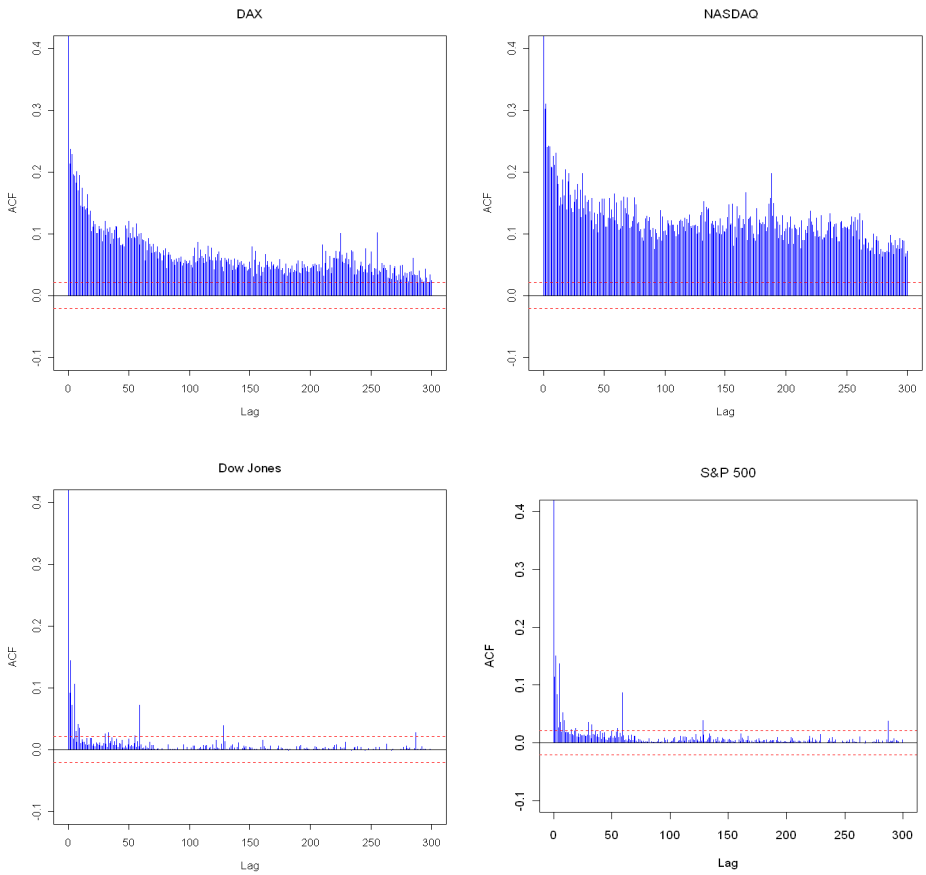


Figure 10: ACFs of squared index returns. Sample period: January 1, 1975 to December 31, 2007.

			DAX		DOW		NASDAQ		S&P 500	
Horizon	Distribution	Model	GARCH	LM	GARCH	LM	GARCH	LM	GARCH	LM
5 days	skewed-$t(3)$	$\overline{VaR}$	6.00	7.79	5.28	5.18	5.71	8.44	4.66	5.02
		$\%\ Viol.$	3.49	1.46	1.80	1.46	1.69	1.46	2.02	1.69
		LR_{uc}	133.78*	1.68	4.64**	1.68	3.52	1.68	7.27*	3.52
			[0.000]	[0.195]	[0.031]	[0.195]	[0.061]	[0.200]	[0.007]	[0.06]
		LR_{ind}	2.54	0.39	9.64*	6.34**	0.51	0.38	0.79	1.33
			[0.111]	[0.535]	[0.002]	[0.012]	[0.473]	[0.535]	[0.373]	[0.249]
		LR_{cc}	36.32*	2.07	14.28**	8.02**	4.03	2.07	8.78*	4.85
			[0.000]	[0.356]	[0.001]	[0.018]	[0.120]	[0.357]	[0.017]	[0.089]
		$\overline{LF}$	2.09	2.16	1.390	1.413	1.80	2.28	1.08	1.08
10 days	skewed-$t(3)$	$\overline{VaR}$	6.02	10.77	5.28	7.22	6.46	10.31	4.75	6.70
		$\%\ Viol.$	10.19	1.56	3.38	1.80	6.53	2.25	4.30	1.80
		LR_{uc}	93.62*	1.27	15.66*	2.33	61.12*	5.19	26.61*	2.33
			[0.000]	[0.260]	[0.000]	[0.127]	[0.000]	[0.023]	[0.000]	[0.127]
		LR_{ind}	0.03	2.87	0.41	0.29	0.01	0.46	0.04	0.30
			[0.871]	[0.090]	[0.523]	[0.588]	[0.935]	[0.500]	[0.834]	[0.588]
		LR_{cc}	93.65*	4.14	16.07*	2.62	61.13*	5.65	26.66*	2.62
			[0.000]	[0.126]	[0.000]	[0.269]	[0.000]	[0.059]	[0.000]	[0.269]
		$\overline{LF}$	2.97	3.53	1.78	2.29	3.39	4.07	1.45	2.28
20 days	skewed-$t(3)$	$\overline{VaR}$	6.08	13.95	4.14	7.47	8.17	15.86	4.14	8.00
		$\%\ Viol.$	14.86	1.80	11.71	1.80	7.66	3.15	5.86	1.35
		LR_{uc}	121.10*	1.16	83.06*	1.17	40.67*	6.62**	24.93*	0.25
			[0.000]	[0.281]	[0.000]	[0.281]	[0.000]	[0.010]	[0.000]	[0.618]
		LR_{ind}	0.32	0.15	0.358	0.15	0.09	0.45	1.62	0.08
			[0.571]	[0.702]	[0.550]	[0.702]	[0.767]	[0.500]	[0.203]	[0.774]
		LR_{cc}	121.40*	1.32	83.42*	1.31	40.76*	7.08**	26.55*	0.33
			[0.000]	[0.519]	[0.000]	[0.519]	[0.000]	[0.030]	[0.000]	[0.847]
		$\overline{LF}$	5.62	5.54	2.20	2.64	6.38	8.28	1.92	3.27
60 days	skewed-$t(3)$	$\overline{VaR}$	6.16	21.03	4.15	10.34	6.14	15.31	3.41	10.43
		$\%\ Viol.$	14.86	2.70	8.11	2.70	16.22	4.05	9.46	1.35
		LR_{uc}	40.37*	1.50	14.98*	1.48	46.17*	3.95**	19.49*	0.08
			[0.000]	[0.224]	[0.000]	[0.224]	[0.000]	[0.047]	[0.000]	[0.773]
		LR_{ind}	0.11	0.11	0.52	0.11	0.00	0.25	0.19	0.03
			[0.743]	[0.739]	[0.472]	[0.739]	[0.963]	[0.615]	[0.665]	[0.869]
		LR_{cc}	40.47*	1.59	15.50*	1.59	46.17*	4.20	19.68*	0.11
			[0.000]	[0.452]	[0.000]	[0.452]	[0.000]	[0.122]	[0.000]	[0.946]
		$\overline{LF}$	31.47	23.24	8.03	8.20	27.96	16.66	9.03	9.05

Table 5: Results of 1%-VaR prediction for various horizons. * and ** denote rejecting the null hypothesis at the 1% and 5% significance level, respectively. The corresponding p-values are given in parenthesis. To save space, we refer to the AR(1)-GARCH(1,1) and AR(1)-GARCH(1,1)-LM model just as "GARCH" and "LM", repectively.

			DAX		DOW		NASDAQ		S&P 500	
Horizon	Distribution	Model	GARCH	LM	GARCH	LM	GARCH	LM	GARCH	LM
5 days	normal	$\overline{VaR}$	2.99	3.88	2.10	2.07	3.01	4.36	2.12	2.26
		$\%\ Viol.$	14.17	9.90	12.93	13.30	14.96	9.67	12.37	11.36
		LR_{uc}	450.20*	252.60*	389.90*	411.60*	489.80*	243.10*	363.30*	316.70*
			[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
		LR_{ind}	9.90*	4.81**	2.16	0.76	1.72	1.80	0.01	0.03
			[0.002]	[0.028]	[0.141]	[0.385]	[0.190]	[0.179]	[0.905]	[0.862]
		LR_{cc}	460.10*	257.40*	392.10*	412.30*	491.50*	244.90*	363.30*	316.70*
			[0.000]	[0.000]	[0.0000	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
		$\overline{LF}$	4.39	3.41	2.96	2.99	5.38	3.95	2.73	2.60
10 days	normal	$\overline{VaR}$	3.00	5.98	2.12	2.80	3.03	6.90	2.15	3.04
		$\%\ Viol.$	18.02	10.01	15.76	11.94	2.93	2.25	16.22	10.59
		LR_{uc}	325.30*	85.58*	263.30*	171.30*	10.98*	5.19**	277.00*	140.90*
			[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.023]	[0.000]	[0.000]
		LR_{ind}	3.01	0.59	1.06	0.02	0.76	0.46	4.43**	0.00
			[0.083]	[0.442]	[0.302]	[0.882]	[0.385]	[0.500]	[0.035]	[0.999]
		LR_{cc}	328.37*	86.17*	266.30*	171.30*	11.74*	5.65	281.50*	140.90*
			[0.000]	[0.000]	[0.000]	[0.000]	[0.003]	[0.059]	[0.000]	[0.000]
		$\overline{LF}$	7.11	3.29	4.41	3.52	2.50	6.00	3.61	2.28
20 days	normal	$\overline{VaR}$	3.02	7.55	2.20	3.73	3.01	6.62	2.13	4.19
		$\%\ Viol.$	23.87	10.36	18.02	13.90	27.03	14.41	18.92	11.26
		LR_{uc}	247.50*	68.02*	162.70*	109.80*	296.80*	115.40*	175.10*	77.96*
			[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
		LR_{ind}	0.73	0.19	0.01	0.03	0.36	0.54	0.18	0.33
			[0.392]	[0.665]	[0.925]	[0.853]	[0.547]	[0.464]	[0.675]	[0.568]
		LR_{cc}	248.20*	68.20*	162.70*	109.90*	297.20*	116.00*	175.30*	78.28*
			[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
		$\overline{LF}$	11.90	4.56	4.83	2.48	16.04	9.18	4.48	1.96
60 days	normal	$\overline{VaR}$	3.07	10.65	2.20	6.40	3.05	12.58	2.13	7.30
		$\%\ Viol.$	22.97	12.16	16.20	6.75	24.32	6.76	16.22	6.76
		LR_{uc}	77.96*	29.42*	46.17*	10.84*	84.80*	10.84*	46.17*	10.84*
			[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.001]	[0.000]	[0.001]
		LR_{ind}	1.77	0.83	0.75	0.73	2.57	0.73	0.00	1.05
			[0.183]	[0.362]	[0.386]	[0.394]	[0.109]	[0.394]	[0.963]	[0.305]
		LR_{cc}	79.73*	30.25*	46.92*	11.56*	87.37*	11.56*	46.17*	11.89*
			[0.000]	[0.000]	[0.000]	[0.003]	[0.000]	[0.003]	[0.000]	[0.003]
		$\overline{LF}$	44.73	28.23	10.89	7.84	43.91	21.05	11.20	7.91

Table 6: Results of 1%-VaR prediction for various horizons. * and ** denote rejecting the null hypothesis at the 1% and 5% significance level, respectively. The corresponding p-values are given in parenthesis. To save space, we refer to the AR(1)-GARCH(1,1) and AR(1)-GARCH(1,1)-LM model just as "GARCH" and "LM", repectively.

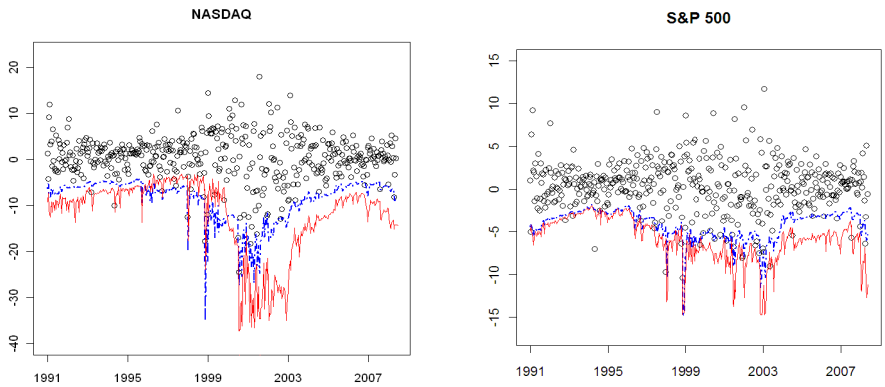


Figure 11: 10-day ahead VaR forecasts for the AR(1)-GARCH(1,1) (blue, dashed line) and AR(1)-GARCH(1,1)-LM (red, solid line) model with skewed student-t innovations from January 1, 1991 to December 31, 2007.

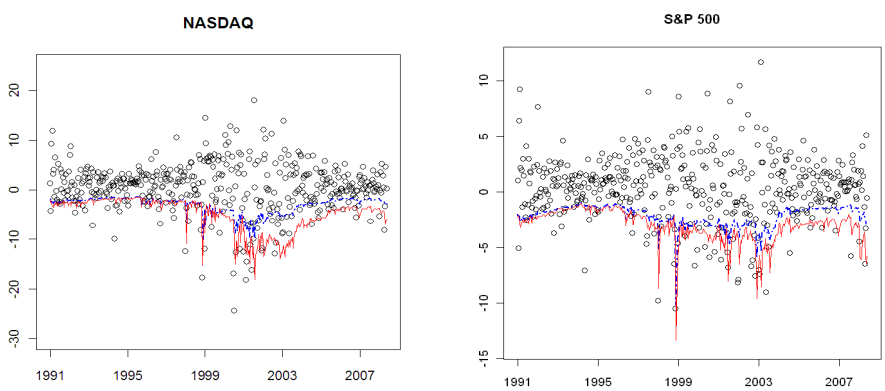


Figure 12: 10-day ahead VaR forecasts for the AR(1)-GARCH(1,1) (blue, dashed line) and AR(1)-GARCH(1,1)-LM (red, solid line) model with normal innovations from January 1, 1991 to December 31, 2007.

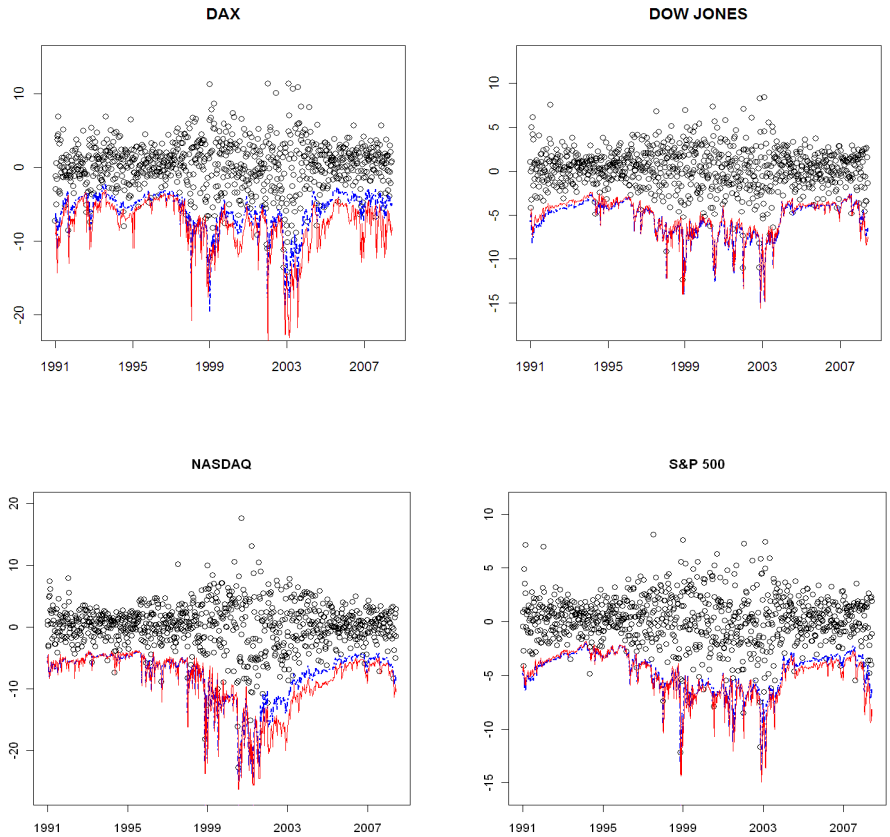


Figure 13: 5-day ahead VaR forecasts for the AR(1)-GARCH(1,1) (blue, dashed line) and AR(1)-GARCH(1,1)-LM (red, solid line) model with skewed student-t distribution from January 1, 1991 to December 31, 2007.

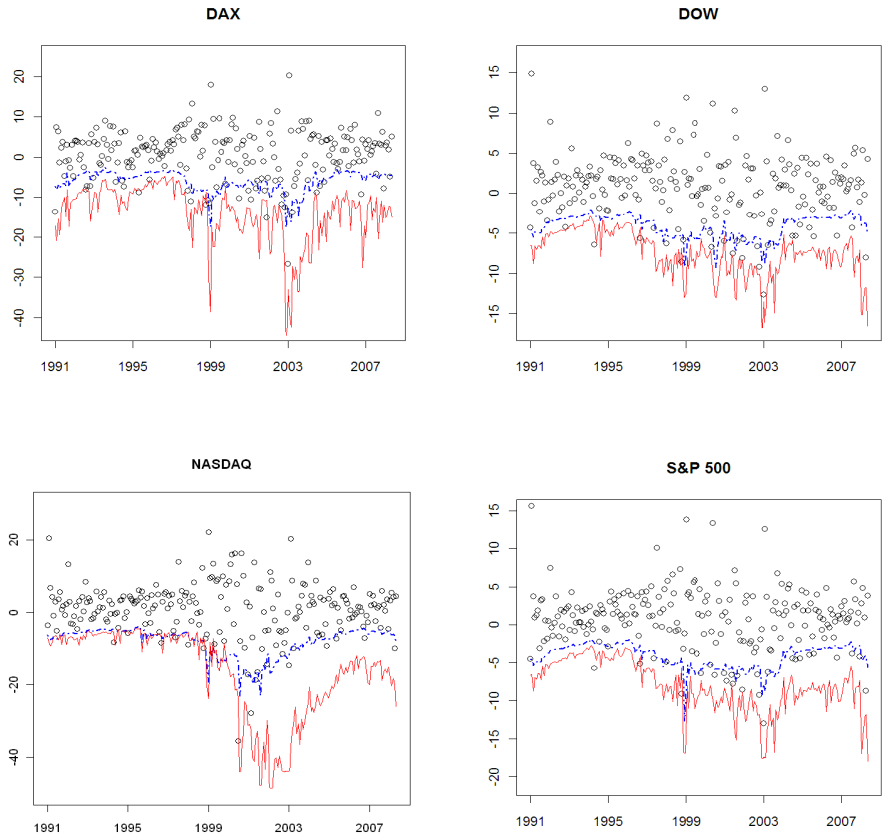


Figure 14: 20-day ahead VaR forecasts for the AR(1)-GARCH(1,1) (blue, dashed line) and AR(1)-GARCH(1,1)-LM (red, solid line) model with skewed student-t innovations from January 1, 1991 to December 31, 2007.

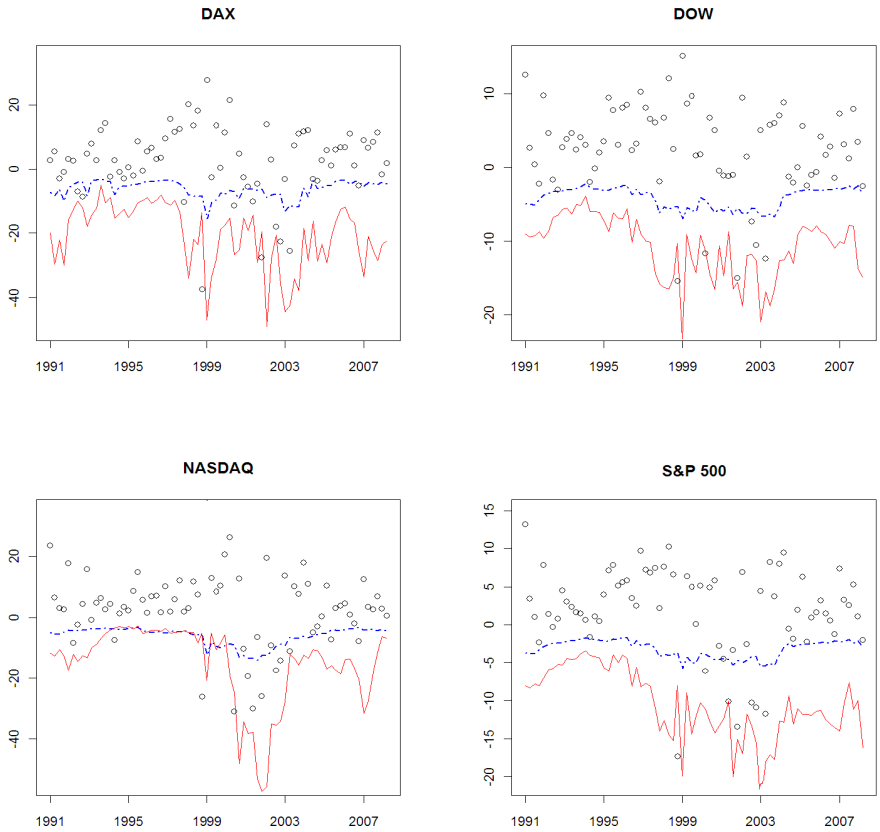


Figure 15: 60-day ahead VaR forecasts for the AR(1)-GARCH(1,1) (blue, dashed line) and AR(1)-GARCH(1,1)-LM (red, solid line) model with skewed student-t innovations from January 1, 1991 to December 31, 2007.

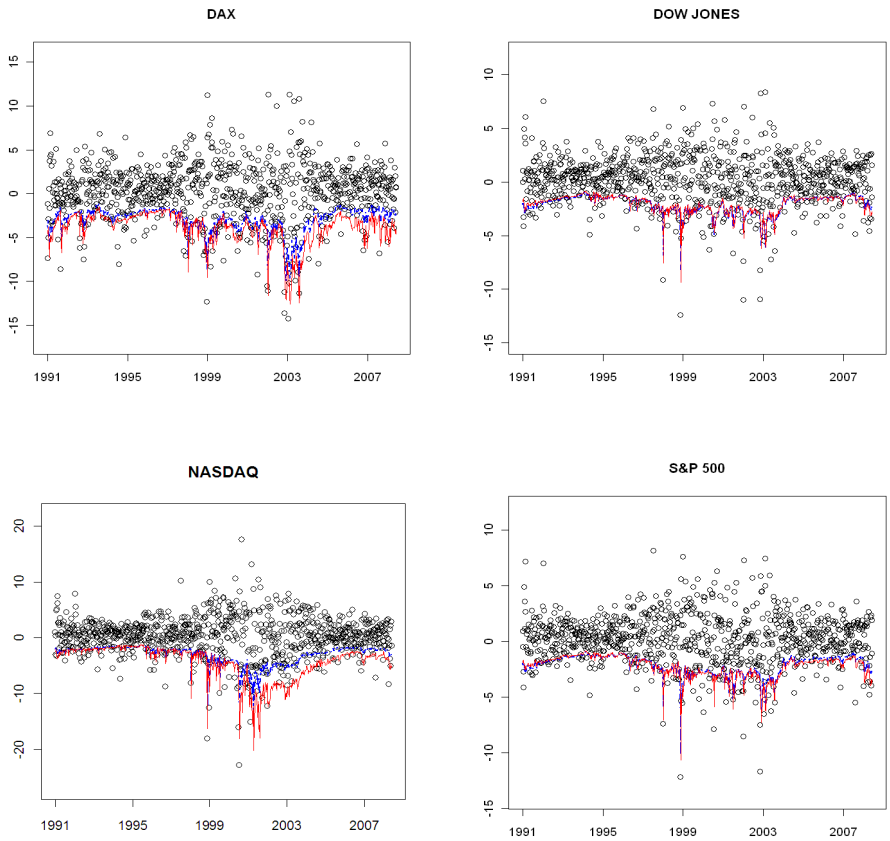


Figure 16: 5-day ahead VaR forecasts for the AR(1)-GARCH(1,1) (blue, dashed line) and AR(1)-GARCH(1,1)-LM (red, solid line) model with normal innovations from January 1, 1991 to December 31, 2007.

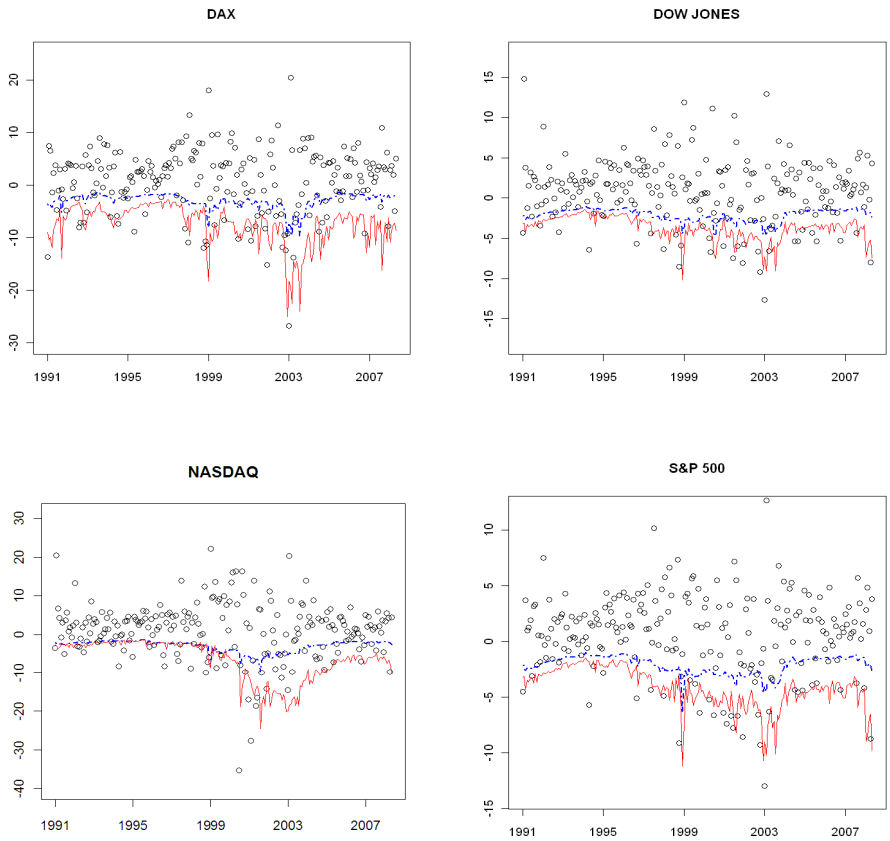


Figure 17: 20-day ahead VaR forecasts for the AR(1)-GARCH(1,1) (blue, dashed line) and AR(1)-GARCH(1,1)-LM (red, solid line) model with normal innovations from January 1, 1991 to December 31, 2007.

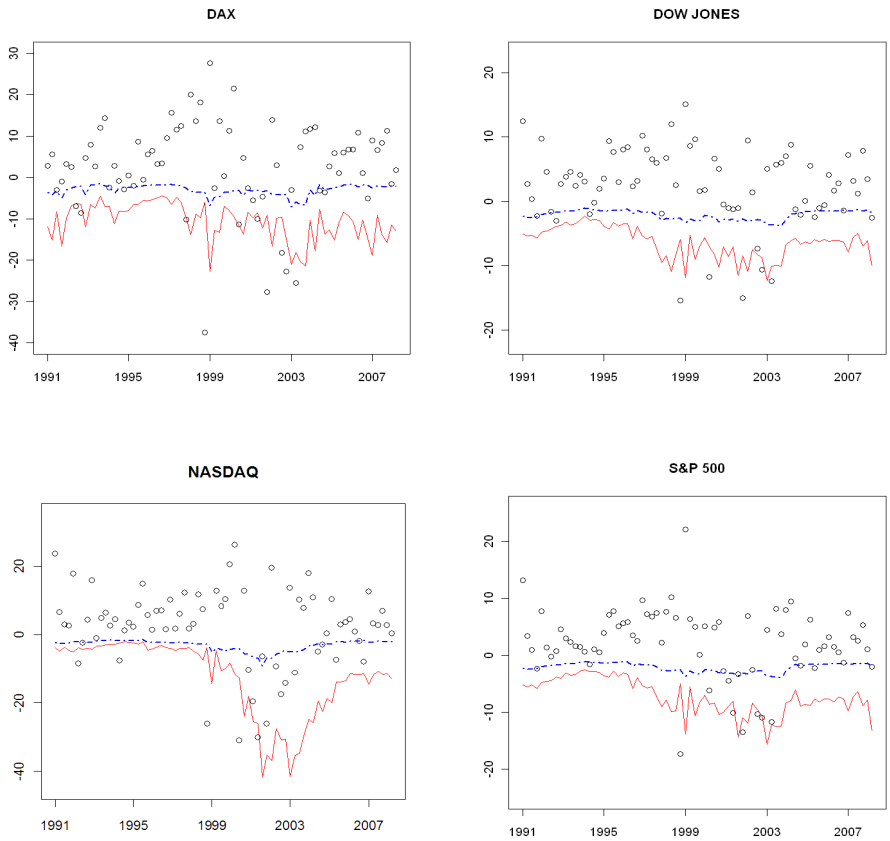


Figure 18: 60-day ahead VaR forecasts for the AR(1)-GARCH(1,1) (blue, dashed line) and AR(1)-GARCH(1,1)-LM (red, solid line) model with normal innovations from January 1, 1991 to December 31, 2007.