Value-at-Risk and Expected Shortfall when there is long range dependence

Wolfgang Härdle*
Julius Mungo*

* Humboldt-Universität zu Berlin, Germany

BERLIN

ECONOMIC RISK

SFB 649

Value-at-Risk and Expected Shortfall when there is long range dependence.

Wolfgang Karl Härdle[*] Julius Mungo[†]

[*]CASE – Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany

[†]CASE – Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany; e-mail: mungo@wiwiw.hu-berlin.de; phone: +49(0)30 2093-5654

Abstract

Empirical studies have shown that a large number of financial asset returns exhibit fat tails and are often characterized by volatility clustering and asymmetry. Also revealed as a stylized fact is Long memory or long range dependence in market volatility, with significant impact on pricing and forecasting of market volatility. The implication is that models that accomodate long memory hold the promise of improved long-run volatility forecast as well as accurate pricing of long-term contracts. On the other hand, recent focus is on whether long memory can affect the measurement of market risk in the context of Value-at-Risk (VaR). In this paper, we evaluate the Value-at-Risk (VaR) and Expected Shortfall (ESF) in financial markets under such conditions. We examine one equity portfolio, the British $FTSE$100 and three stocks of the German DAX index portfolio (Bayer, Siemens and Volkswagen). Classical VaR estimation methodology such as exponential moving average (EMA) as well as extension to cases where long memory is an inherent characteristics of the system are investigated. In particular, we estimate two long memory models, the Fractional Integrated Asymmetric Power-$ARCH$ and the Hyperbolic-$GARCH$ with different error distribution assumptions. Our results show that models that account for asymmetries in the volatility specifications as well as fractional integrated parametrization of the volatility process, perform better in predicting the one-step as well as five-step ahead VaR and ESF for short and long positions than short memory models. This suggests that for proper risk valuation of options, the degree of persistence should be investigated and appropriate models that incorporate the existence of such characteristic be taken into account.

2

1 Introduction

Empirical studies have shown that a large number of financial asset returns exhibit fat tails (leptokurtosis) and are often characterized by volatility clustering (large changes that tend to be followed by large changes and vice versa), leverage effects (changes in stock prices that tend to be negatively correlated with changes in volatility) as well as long range dependence (signs of dependence between far apart observations in a given time series). Ding et al. (1993), So (2000), Henry (2002), Aloui et al. (2005) have found that financial assets may exhibit long range dependence on stock market volatility. Bollerslev and Mikkelsen (1996), Herzberg and Sibbertsen (2004) have shown that such long term dependence has significant impact on the pricing of financial derivatives as well as forecasting market volatility. This may in turn affect an investment portfolio, bringing about huge losses or gains.

An important part of modern financial risk management is the measuring and forecasting portfolio risk and fluctuations due to changing market and/ or portfolio shares. A number of applications of Value-at-risk (VaR) on financial asset returns have shown that due to excess kurtosis, estimation under the assumption of normally distributed residuals tend to underestimate or overestimate the true VaR. Brooks and Persand (2003a) showed that models that do not allow for asymmetries in the unconditional distribution or volatility specification, underestimate the true VaR and are most likely to generate inaccurate forecasts. Giot and Laurent (2004) pointed out that the VaR for daily stock indexes with a skewed student-t distribution performed better than with pure symmetric distribution.

On the other hand, studies on whether long memory phenomena can affect the measurement of market risk in the context of value-at-risk (VaR) is a topic of interest. In this paper, we focus on financial market risk which refers to the risk that arises due to unforeseen movements in markets price (e.g. equity or option prices) or market rates (e.g. interest or exchange rates). We examine one equity portfolio, the British $FTSE100$ and three stocks of the German DAX index portfolio (Bayer, Siemens and Volkswagen). In addition to applying classical risk estimation methodology such as the exponential moving average (EMA) and the empirical distribution function (EDF), our
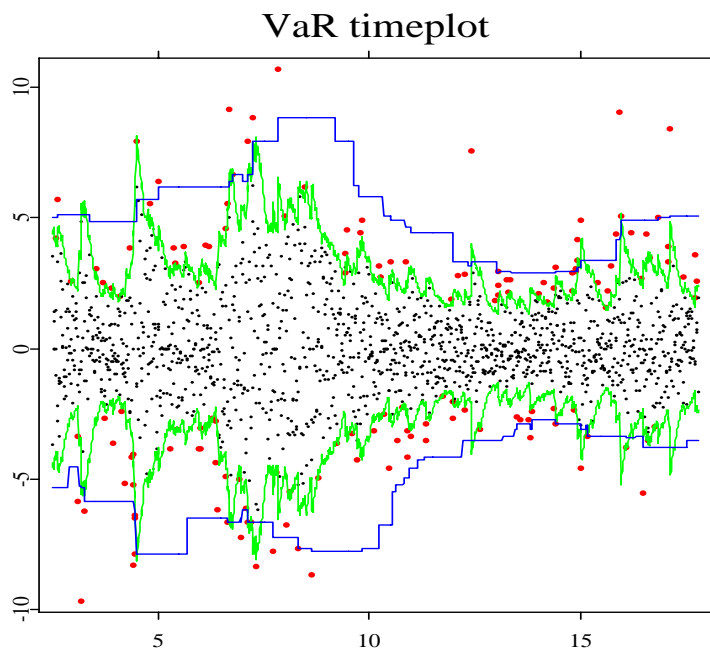
VaR timeplot

Figure 1: *Time plots of VaR forecasts at 99% and 1% level using (EDF) and (EMA) (with decay factor, $\lambda = 0.94$ and a 250 days window) for the Volkswagen stock returns from $3rd$ January 2000 - $10th$ October 2006. Returns which exceed the VaR are marked in red. A total of 1781 data points.*

focus is on risk evaluation in the case where long memory is an inherent characteristics of the system. Figure 1 shows a time plot of VaR forecasts using the exponential moving average (EMA) method (with decay factor, $\lambda = 0.94$ and a rolling window of 250 days) and the empirical distribution function (EDF) method on Volkswagen log-returns from $3rd$ January 2000 - $10th$ October 2006. The significance level under consideration is 1% for large losses and 99% for large profits. Returns which exceed the VaR are marked in red. Though the EMA method shows better forecast result than the (EDF), the poor forecast quality of these methods are reflected by the number of VaR exceeded. This indicates a possible inadequacy of an assumed normal distribution. Guermat and Harris (2002) have shown that the EMA-based VaR forecasts are excessively volatile and unnecessarily high, when returns are not conditionally normal.

In what follows, we focus on risk evaluation under long range depen-

dence in the volatility process. In particular, we apply two long memory in volatility models, the Fractional Integrated Asymmetric Power-$ARCH$ ($FIAPARCH$) and the Hyperbolic-$GARCH$ ($HYGARCH$) with different error distribution assumptions. Our results show that models that account for asymmetries in the volatility specifications as well as fractional integrated parametrization of the volatility process, perform better in predicting the one-step as well as five-step ahead VaR and ESF for short and long positions than short memory models. For example, results for the DAX stocks suggest that accurate VaR estimates may be obtained from a long memory in volatility model, coupled with skewed Student-t distributed innovations. This indicates that for proper risk valuation of these stocks, the degree of persistence should be investigated and appropriate models that incorporate the existence of such characteristic be taken into account.

Measuring financial market risk can be described by the VaR, which estimates the probability of a portfolio of assets losing a specific amount over a specified time period due to adverse movements in the underlying market factors, Jorion (2001).

Let $S_1, \ldots, S_n$ be a time series of stock's prices. For the price series S_t ($t = 1, \ldots, n$), daily logarithmic return (log-return) is defined as $r_t = \log\left(\frac{S_t}{S_{t-1}}\right)$. The loss that r_t could incur over a given time horizon h with a given confidence level $1 - \alpha$, at time t conditionally on available information, $\mathcal{F}_{t-1}$ is defined by the VaR:

$$P\left\{(r_t \geq VaR_{t,h}(\alpha)|\mathcal{F}_{t-1}\right\} = \alpha.$$

VaR can also be defined from the distribution of log-returns, Jorion (2001). If q_α is the αth percentile of continuously compounded return r_t, then the $h-$step ahead VaR in time t is defined as

$$VaR_{t,h}(\alpha) = (1 - e^{q_\alpha})S_{t-h}.$$

This definition shows that a good VaR estimate can be realized with accurate forecasts of the percentiles q_α, which depends on appropriate volatility modelling. VaR can describe extreme events, but it can not aggregate risk

in the sense of subadditivity on portfolios (in this case, risk is diversified). This recognized shortcoming is addressed by the concept of expected shortfall (ESF), a "coherent risk measure" in the sense of Artzner et al. (1999).

The ESF is a coherent measure for such risk, used to predict the expected value of the loss given that the loss is greater than the VaR, Scaillet (2004).

$$ESF_t = \mathrm{E}(|L_t| > |VaR_t|),$$

where L_t is the expected value of loss if a VaR_t violation occurs.
An equivalent measure is the average multiple of tail event to risk ($AMTERM$), that "measures how much events in tail distribution exceed the VaR, by calculating the average multiple of these outcomes to their corresponding VaR measures", Hendricks (1996).

In our application, we estimate simultaneously VaR and ESF for the long and short positions. We apply long memory in volatility models using Student-t and skewed student-t distribution assumptions to find out which model has the best overall performance. Without including all long memory GARCH-type specification models available in the literature, we estimate only models that are extensions of specifications implemented in similar studies. We employ evaluation by backtesting procedure to investigate and select models that predict the VaR accurately and minimizes, if a VaR violation occurs.

In this regard, using daily returns, we analyze a VaR and ESF for daily trading portfolios made up of long, (VaR_L) and short, (VaR_S) positions, using various long-memory GARCH-type models. At long position, traders incur losses when drop in stock prices (negative returns) are observed as opposed to short position where losses are incurred when increase in stock prices (positive returns) are observed.

Our analysis is twofold; first, we test and estimate long range dependence in the return indexes by employing the rescaled variance test, Geweke and Porter-Hudak (1983) (GPH) and two estimators, the log-periodogram regression (GPH), Geweke and Porter-Hudak (1983) and the Gaussian semi-parametric (GSP), Robinson (1995a) estimators. We then model long range dependence by fitting GARCH-type models that are known to accommo-

date distributional characteristics such as volatility clustering, asymmetry, leptokurtosis as well as long memory. This includes the $(HYGARCH)$, Davidson (2004) model that can well accommodate long memory and the $(FIAPARCH)$, Tse (1998) that can accommodate both asymmetry and long memory.

Second, we assume that investors in today's financial market can invest by holding not only long but short positions. We compute the VaR and ESF for both positions based on the chosen models. Our results suggest that accurate VaR estimates may be obtained from a long memory in volatility model, coupled with skewed Student-t distributed innovations. Although these models can achieve accurate VaR and ESF for insample one-step ahead prediction, we find that models that account for asymmetries as well as fractional integrated parametrization of the volatility process, perform better than those that reflect only symmetry or long-memory. This confirms, the findings of Angelidis and Degiannakis (2006), in suggesting that models with fractional integration parametrization of the volatility process are necessary for accurate assessment of market risk. Long- memory in volatility models hold the promise of improved long-run volatility forecast and more accurate pricing of long-term contracts. The implication can therefore be extended to the context of risk management.

An outline of the remaining sections of this paper is as follows. In section 2 we characterize the long memory models used and their specifications. Section 3 present and overview on Backtesting VaR valuation methods. In section 4 we present our empirical application and results.

2   Long Memory Models Specifications

We consider the continuously compounded rate of returns of a stock index, $r_t = \log\left(\frac{S_t}{S_{t-1}}\right)$ (S_t is the stock price at time t) from time $t-1$ to t, which can be decomposed into a predictable component, μ_t and an unpredictable

component, ε_t, that:

$$r_t = \mu_t + \varepsilon_t, \qquad (1)$$

$$\varepsilon_t = z_t \sigma_t \qquad (2)$$

Following Bollerslev (1986), the $GARCH(p,q)$ model for conditional variance is specified as

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \varepsilon_{t-j}^2 \qquad (3)$$

where $\omega > 0$, $\alpha_i \geq 0$ for $i = 1, \ldots, q$ and $\beta_j \geq 0$ for $j = 1, \ldots, p$. In the case where $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$, the process ε_t is covariance stationary with unconditional variance

$$\sigma^2 = \frac{\alpha}{1 - \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j}.$$

A shock to σ_t^2 is known to have decaying impact on future volatility, σ_{t+h}^2, for increasing h, but is asymptotically negligible. Many high-frequency times series applications that have employed $GARCH(p,q)$ process to estimate σ_t^2, show strong dependence in that $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j \approx 1$.

In the framework of risk management, the application of the $GARCH(1,1)$ model to forecast investors risk have yield mixed results. Billo and Pelizzon (2000) have shown that for VaR valuation, the number of exceptions generated by the $GARCH(1,1)$ model deviates significantly from the theoretical values. Using lag-operator L, equation (3) can be expressed as an $ARMA$ process:

$$[1 - \alpha(L) - \beta(L)]\varepsilon_t^2 = \omega + [1 - \beta(L)](\varepsilon_t^2 - \sigma_t^2)$$

where $\alpha(L) = \alpha_1(L) + \cdots + \alpha_q L^q$ and $\beta(L) = \beta_1(L) + \cdots + \beta_p L^p$.

For $\alpha(L) + \beta(L) = 1$, Engel and Bollerslev (1986) proposed the Integrated-$GARCH$, $IGARCH(p,q)$ model, expressed as

$$\phi(L)(1 - L)\varepsilon_t^2 = \omega + [1 - \beta(L)](\varepsilon_t^2 - \sigma_t^2) \qquad (4)$$

where $\phi(L) = [1 - \alpha(L) - \beta(L)](1 - L)^{-1}$, is of order $max(p,q) - 1$. Compared

to $GARCH(p,q)$, the σ_t^2 in $IGARCH(p,q)$ is infinite and a shock to σ_t^2 is persistent, indicating its importance for volatility forecasting.

A special case of the $IGARCH(p,q)$ is the **Exponential Weighted Moving Average** ($EWMA$) model or RiskMetrics, expressed as

$$\sigma_t^2 = (1-\lambda)\varepsilon_{t-1}^2 + \lambda\sigma_{t-1}^2 \tag{5}$$

where $\lambda = 0.94$ for daily and $\lambda = 0.97$ for weekly data respectively. In this case the $ARCH$ and $GARCH$ coefficients are fixed. This model process provides a cumulative impulse response function that is constant over time, which is not very reliable. So and Yu (2006) have shown that for the $99\%\ VaR$ valuation, the $EWMA$ model is outperformed by stationary and fractionally integrated $GARCH$ models.

Between stationarity and unit root, the concept of fractional integration has been extended to $GARCH$ type models, with many ways to parametrisize the conditional volatility. In our application, we consider two long memory in volatility models, the $FIAPARCH$ of Tse (1998) and the $HYGARCH$ of Davidson (2004).

FIAPARCH Model

Tse (1998) introduced this model by incorporating a fractional filter $(1-L)^d$ into the Asymmetric Power $ARCH$ ($APARCH$) model frame work of Ding et al. (1993). The $APARCH(p,q)$ can be expressed as

$$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i(|\varepsilon_{t-i}| - \gamma_i\varepsilon_{t-i})^\delta + \sum_{j=1}^p \beta_j\sigma_{t-j}^\delta, \tag{6}$$

where $\delta > 0$ and $-1 < \gamma_i < 1$. δ represent a Box-Cox transformation of σ_t and γ_i the leverage effect. For the $FIAPARCH(p,d,q)$ model of Tse (1998), the conditional volatility specification is given as

$$\sigma_t^\delta = \omega + \left\{1 - [1 - \beta(L)]^{-1}\alpha(L)(1-L)^d\right\}(|\varepsilon_t| - \gamma\varepsilon_t)^\delta. \tag{7}$$

$(1 - L)^d$ is the fractional filter defined by its binomial expansion

$$(1 - L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j - d)}{\Gamma(-d)\Gamma(j + 1)} L^j \tag{8}$$

where $\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$ is the gamma function. The autocorrelation function is defined by

$$\rho_k = \frac{\Gamma(1 - d)\Gamma(k + d)}{\Gamma(d)\Gamma(k + 1 - d)} \sim Ck^{2d-1}$$

If k is large, the coefficients in equation (8) for $(0 < d < 1)$ decay hyperbolically. Therefore the effect of past innovations on current conditional variance dies out at a hyperbolic rate with increasing lag length.

The term $(|\varepsilon_t| - \gamma\varepsilon_t)^\delta$ reflects "leverage" effect, $\gamma \neq 0$. When $0 < \gamma < 1$, a positive innovation increases volatility less than a negative innovation and vice versa for $-1 < \gamma < 0$. For $\gamma = 0$, positive and negative innovations of same magnitude have thesame effect on volatility. The $FIAPARCH$ model allows for volatility clustering, leptokurtosis, long memory and features asymmetry in relationship between the conditional variance and the lagged squared innovations. It nests the $GARCH$, $IGARCH$, $GRJ - GARCH$, $APARCH$, $FIPARCH$, $FIGARCH$ and $FIAGARCH$ models as special cases.

HYGARCH Model

To test whether nonstationarity of the fractional integrated GARCH process $FIGARCH(p, d, q)$ of Baillie et al. (1996) holds, Davidson (2004) proposed the hyperbolic GARCH ($HYGARCH$) model. The $FIGARCH$ process is expressed as

$$\Phi(L)(1 - L)^d \varepsilon_t^2 = \omega + \Theta(L)\nu_t \tag{9}$$

with $\nu_t = \varepsilon_t^2 - \sigma_t^2$. The conditional variance for the $FIGARCH(p, d, q)$ can

be represented as

$$\sigma_t^2 = \frac{\omega}{1-\theta(L)} + \left\{1 - \frac{\phi(L)(1-L)^d}{1-\theta(L)}\right\}\varepsilon_t^2 \qquad (10)$$

$$= \omega[1-\theta(L)]^{-1} + \left\{1 - [1-\theta(L)]^{-1}\phi(L)(1-L)^d\right\}\varepsilon_t^2 \qquad (11)$$

where $0 \le d \le 1$.

The $HYGARCH$ model is obtained by extending the conditional variance of the $FIGARCH(p,d,q)$ by introducing weights to its difference operator. By replacing $1 - [1 - \theta(L)]^{-1}\phi(L)(1-L)^d$ in equation (11) by $1 - [1 - \theta(L)]^{-1}\phi(L)\left[1 + \alpha\left\{(1-L)^d\right\}\right]$, the conditional variance of the $HYGARCH$ process can be formulated as

$$\sigma_t^2 = \omega[1-\theta(L)]^{-1} + \left\{1 - [1-\theta(L)]^{-1}\phi(L)\left[1 + \alpha\left\{(1-L)^d\right\}\right]\right\}\varepsilon_t^2 \quad (12)$$

This model is able to reproduce the characteristics of volatility clustering, leptokurtosis and long memory but disregards asymmetry and the fact that the conditional volatility is best represented by non-integer powers of the absolute value of the observations. It nests the $GARCH$ models (for $\alpha = 0$), $IGARCH$ (for $\alpha = d = 1$) and $FIGARCH$ (for $\alpha = 1$ or $\log \alpha = 0$).

2.1 Long Memory Test and Estimation

Many empirical studies have focused on testing the presence of long memory in volatility of financial markets. Motivations for such studies include the use of volatility as preliminary measure of risk in financial markets and their predictions can lead to improve forecasts of asset price movement. In this paper, long memory in volatility in an equity portfolio (British $FTSE$) and German DAX stocks (Bayer, Siemens, Volkswagen) is examined using two proxies of volatility returns: absolute and square returns. The rescaled variance (V/S) test of Giraitis et al. (2003) and two semiparametric estimators of long memory parameter, the log-periodogram regression (GPH) of Geweke and Porter-Hudak (1983) and the Gaussian semiparametric (GSP) of Robinson (1995a) are employed.

Rescaled Variance test

Giraitis et al. (2003) introduced the Rescaled Variance test (V/S), by centering the Kwiatkowski et al. (1992), $(KPSS)$ test of trend stationarity against a unit root alternative, based on the partial sum of deviations from the mean. For testing long memory in a stationary series the $KPSS$ statistic is given as

$$KPSS_T = \frac{1}{T^2 \hat{\sigma}_T^2(q)} \sum_{k=1}^{T} \left\{ \sum_{j=1}^{k} (r_j - \bar{r}_T) \right\}^2 \tag{13}$$

where $\widehat{\sigma}_T^2(q) = \hat{\gamma}_0 + 2 \sum_{j=1}^{q} \left(1 - \frac{j}{1+q}\right) \hat{\gamma}_j$, is the heteroscedastic and auto-correlation consistent (HAC) estimator of the variance, $(q < T)$. $\hat{\gamma}_0$ is the variance of the process and the sequence $\{\hat{\gamma}_j\}_{j=1}^{q}$ denotes the autocovariances of the process up to the order q.

Centering $KPSS_T$ based on the partial sum of the deviations from the mean, the Rescaled Variance test statistic is expressed as:

$$V/S(q) = \frac{1}{T^2 \hat{\sigma}_T^2(q)} \left[\sum_{k=1}^{T} \left\{ \sum_{j=1}^{k} (r_j - \bar{r}_T) \right\}^2 - \frac{1}{T} \left\{ \sum_{k=1}^{T} \sum_{j=1}^{k} (r_j - \bar{r}_T) \right\}^2 \right] \tag{14}$$

Giraitis, Kokoszka, Leipus (2001) have shown that this statistic can detect long range dependence in the volatility for the class of $ARCH(\infty)$ processes.

Log-Periodogram Estimator

Geweke and Porter-Hudak (1983) (GPH), proposed a semi-parametric procedure to obtain an estimate of the fractional difference parameter based on the periodogram of a time series z_t, $(t = 1, \ldots, T)$ defined by

$$I(\lambda_j) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} z_t e^{-\mathbf{i}\lambda t} \right|^2 \tag{15}$$

where $\lambda_j = \frac{2\pi j}{T}$, $j = 1, \ldots, m$ (m is a positive integer). The memory parameter $\widehat{d}_{GPH}$ is the OLS estimator in a linear regression of the log-periodogram,

$\log I(\lambda_j)$ of the process on a constant and a variable $X_j = \log\left\{4\sin^2(\lambda_j/2)\right\}$:

$$\widehat{d}_{GPH} = -\frac{\sum_{j=1}^{m}(X_j - \bar{X})\log\left\{I(\lambda_j)\right\}}{2\sum_{j=1}^{m}(X_j - \bar{X})} \tag{16}$$

The theoretical asymptotic variance of the spectral regression error term is known to be $\frac{\pi^2}{6}$. Details on the consistency and asymptotic normality of the GPH estimator can be found in Geweke and Porter-Hudak (1983), Robinson (19990), Hassler (1993a). For nonstationary fractional integrated process, Kim and Phillips (2006) have shown that this estimator is consistent in the range $0.5 < d \le 1$.

Gaussian Semiparametric Estimator

The spectrum of a stationary process with long memory parameter d can be approximated in the neighborhood of the zero frequency

$$f(\lambda) \sim C\lambda^{-2d} \tag{17}$$

as $\lambda \to 0+$, $C > 0$. The GSP estimator is derived from the approximation, $\lim_{\lambda_i \to 0+} f(\lambda_i) = C\lambda_i^{-2d}$ of a long memory process in the Whittle approximate maximum likelihood estimator, $L_W(\theta)$. For $m^* = [\frac{T}{2}]$, an approximation to the Gaussian likelihood, Beran (1994) is given as

$$L_W(\theta) = -\frac{1}{2\pi}\sum_{j=1}^{m^*}\log f_\theta(\lambda_j) + \frac{I_T(\lambda_j)}{f_\theta(\lambda_j)} \tag{18}$$

for a given parametric spectral density $f_\theta(\lambda)$. d is estimated by solving the minimization,

$$\operatorname*{arg\,min}_{C,d} L(C,d) = \frac{1}{m}\sum_{j=1}^{m}\left\{\log(C\lambda_j^{-2d}) + \frac{I(\lambda_j)}{C\lambda_j^{-2d}}\right\}$$

where $I(\lambda_j)$ is the periodogram evaluated for a degenerated range of m harmonic frequencies, $\lambda_j = \frac{2\pi j}{T}$, $(j = 1,\dots,m << [\frac{T}{2}])$. [.] represents the integer part operator, bounded by the bandwidth parameter m, which increases with the sample size T but more slowly. The bandwidth m must

satisfy $\frac{1}{m} + \frac{m}{T} \to 0$ as $T \to \infty$. For $m = [\frac{T}{2}]$, this estimator is Gaussian for the parametric model $f(\lambda) = C\lambda^{-2d}$.

The Gaussian Semiparametric Estimator estimator for d, $\hat{d}_{GSP}$ is expressed as

$$\hat{d}_{GSP} = \arg\min_d \left\{ \log\left(\frac{1}{m} \sum_{j=1}^{m} \frac{I(\lambda_j)}{C\lambda_j^{-2d}} \right) - \frac{2d}{m} \sum_{j=1}^{m} \log(\lambda_j) \right\}. \qquad (19)$$

Robinson (1995a) showed that $\sqrt{m}(\hat{d}_{GSP} - d) \xrightarrow{d} N(0, 1/4)$ and Robinson (1999) showed its validity in the presence of some form of conditional heteroscedasticity. In general, the ranges of consistency and asymptotic normality for GSP are the same as those of the GPH estimator, Phillips and Shimotsu (2004).

3 Value-at-Risk Analysis

Mathematically, VaR at α probability level can be defined as

$$VaR_{\alpha,t} = F_t^{-1}(\alpha) = \sigma_t q_\alpha$$

where F_t^{-1} is the inverse function of the conditional cumulative distribution of the underlying at time t, σ_t is the volatility. In this case, the accuracy of the VaR forecast depends heavily on proper estimation of the volatility σ_t and the assumption of the underlying distribution. In our application, we characterize the $FIAPARCH$ and $HYGARCH$ models under the student-t and skewed student-t distribution of the error term and compare their performance.

Suppose that $f(z_t)$ is the *pdf* of the standardized residuals z_t. We characterize different specifications for $f(z_t)$, namely student-t, $z_t \sim st(0, 1; \nu)$ and skewed student-t, $z_t \sim skst(0, 1; \xi, \nu)$ where ν is the degrees of freedom and ξ the asymmetry parameter.

14

The **Student-t** distribution for $z_t \in \mathbb{R}$ is given by

$$f_{st}(z_t; \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu(\nu-2)}} \left(1 - \frac{z_t^2}{\nu-2}\right)^{-\frac{\nu+1}{2}} \tag{20}$$

where the degree of freedom ν, $(\nu > 2)$ and $\Gamma(\nu) = \int_0^\infty e^{-x} x^{\nu-1} dx$ is the gamma function. Under the Student-t distribution, the VaR of α quantile for the long and short position is expressed as

$$VaR_L = \hat{\mu}_t + st_{\alpha,\nu} \hat{\sigma}_t \tag{21}$$

$$VaR_S = \hat{\mu}_t + st_{1-\alpha,\nu} \hat{\sigma}_t \tag{22}$$

where $st_{\alpha,\nu}$ and $st_{1-\alpha,\nu}$ are the left and right quantiles respectively at $\alpha\%$ for the (standardized) student-t distribution and $\hat{\mu}_t$, estimate of the sample mean. The student-t distribution has thicker tails and can deal with leptokurtosis, Bollerslev (1987). For independent and student-t distributed innovations, the conditional variance and kurtosis may vary freely over time, Brooks et al. (2005).

The **skewed student-t** distribution has a density for $z_t \in \mathbb{R}$:

$$f_{skst}(z_t; \nu, \xi) = \begin{cases} \frac{2}{\xi + \frac{1}{\xi}} sg\left\{\xi(sz_t + m); \nu\right\} & \text{if} \quad z_t < -\frac{m}{s} \\ \frac{2}{\xi + \frac{1}{\xi}} sg\left\{\frac{(sz_t + m)}{\xi}; \nu\right\} & \text{if} \quad z_t \geq -\frac{m}{s}, \end{cases} \tag{23}$$

where $g(.; \nu)$ is the symmetric (unit variance) student density, ξ the asymmetric coefficient, ν the degree of freedom. m and s^2 are the mean and variance respectively of the non-standardized $skst$, Lambert and Laurent (2001).

$$m = \frac{\Gamma(\frac{\nu+1}{2})\sqrt{\nu-2}}{\sqrt{\pi}\Gamma(\frac{\nu}{2})} \left(\xi - \frac{1}{\xi}\right),$$

$$s = \sqrt{\left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2}.$$

The skewed student-t distribution is able to reduce excess kurtosis and skew-

ness in standardized residuals, providing efficiency gains. It nests the student-t for $\xi = 1$ and the normal distribution for $\xi = 1$ and $\nu = \infty$.

Under skewed student-t distribution the VaR of α quantile for the long and short position is expressed as

$$VaR_L = \hat{\mu}_t + skst_{\alpha,\nu,\xi}\hat{\sigma}_t \qquad (24)$$
$$VaR_S = \hat{\mu}_t + skst_{1-\alpha,\nu,\xi}\hat{\sigma}_t \qquad (25)$$

where $skst_{\alpha,\nu,\xi}$ and $skst_{1-\alpha,\nu,\xi}$ are the left and right quantiles respectively at $\alpha\%$.

3.1 Backtesting VaR

To compare the accuracy of the different VaR statistical forecasts, tests based on Kupiec (1995) and Christoffersen (1998) backtesting measures are applied.

Unconditional coverage

The basic idea of the Kupiec (1995) test is to examine whether the failure rate of a model is statistically equal to the expected one (unconditional coverage). Let $N = \sum_{t=1}^{T} I_t$ be the number of exceptions in the sample of size T, where

$$I_{t+1} = \begin{cases} 1 & \text{if} \quad r_{t+1} < VaR_{t+1|t}(p) \\ 0 & \text{if} \quad r_{t+1} \geq VaR_{t+1|t}(p), \end{cases}$$

follow a binomial distribution, $N \sim B(T,p)$. If $p = \mathrm{E}[\frac{N}{T}]$ is the expected exception frequency (expected ratio of violations), then the hypothesis for testing whether the failure rate of the model equals the expected one, is given as

$$H_0 : p = p_0 \qquad (26)$$

where p_0 is a pre-specified VaR level. The appropriate likelihood ratio statistic under H_0 is expressed as

$$LR_1 = -2\log\left\{p_0^N(1-p_0)^{T-N}\right\} + 2\log\left\{(N/T)^N(1-N/T)^{T-N}\right\} \qquad (27)$$

Asymptotically, $LR_1 \sim \chi^2(1)$ distributed under H_0. A preferred model for VaR prediction should exhibit the property that the unconditional coverage, measured by $p = \mathrm{E}[\frac{N}{T}]$, equals the desired coverage level, p_0.

Conditional coverage

The Christoffersen (1998) likelihood ratio statistics, test the assumption of unconditional coverage and independence of failures. Let $\pi_{ij} = P(1_t = j | 1_{t-1} = i)$ be the transition probability and $n_{ij} = \sum_{t=1}^{T} 1(1_t = j \ and \ 1_{t-1} = i)$, where $i, j = 0$ or 1. The test hypothesis of independence is given as:

$$H_0 : \pi_{00} = \pi_{10} = \pi, \quad \pi_{01} = \pi_{11} = 1 - \pi \tag{28}$$

One can test the hypothesis of unconditional coverage and independence using the likelihood ratio statistic:

$$LR_2 = -2\log\left\{\hat{\pi}^{n_0}(1 - \hat{\pi})^{n_1}\right\} + 2\log\left\{\hat{\pi}_{00}^{n_{00}}\hat{\pi}_{01}^{n_{01}}\hat{\pi}_{10}^{n_{10}}\hat{\pi}_{11}^{n_{11}}\right\} \tag{29}$$

where $\hat{\pi}_{ij} = \frac{n_{ij}}{(n_{ij}+n_{i,1-j})}$, $n_j = n_{0j} + n_{1j}$ and $\hat{\pi} = \frac{n_0}{(n_0+n_1)}$. Under H_0, $LR_2 \sim \chi^2(2)$. This test has an advantage in that it can reject a VaR model that generates either too many or too few clustered violations, i.e. it takes into account the conditionality in the forecast and enables one to separate clustering effects from distributional assumption effects.

4   Empirical Application

The data consist of daily closing prices for the English $FTSE$ stock index $(04.01.2000 - 30.10.2006)$ and three German DAX stocks; the Bayer, Siemens and Volkswagen, each from $(04.01.2000 - 30.10.2006)$. For each series, the log-returns is expressed (in %) as,

$$r_t = 100 * \log\left(\frac{S_t}{S_{t-1}}\right)$$

denotes the continuously compounded daily returns where S_t is the stock price at time t. Log-returns have the nice property in that the $T - day$

log-returns are the arithmetic sums of one-day log-returns and therefore one can relate between one-day and T-day Value-at-Risk. Summary statistics presented in Table 1 show that all returns share similar properties in kurtosis. The $FTSE$ index and the Volkswagen stock are negatively skewed. The Box-Pierce statistic for remaining serial correlation in the squared standardized residuals (using 24 lags) indicates high serial correlation in the variance. From the descriptive graphics presented in Figure 3, several volatility periods can be observed, such as the periods ending 2001, 2002 and the beginning of 2003. Figure 4 shows the boxplots of log-returns for the four data sets. It also provides evidence of changes in volatilty from time to time.

The V/S test results for long range dependence in the absolute and square returns are presented in Table 2. At 5% level, we fail to reject long memory since the evaluated statistic is over the critical value, 0.1869. The GPH estimates (bandwidths, $m = T^{0.5}, m = T^{0.6}, m = T^{0.8}$) and the GSP estimates (bandwidths, $m = [\frac{T}{4}]$, $m = [\frac{T}{16}]$, $m = [\frac{T}{64}]$) are reported in Table 3. Results from both estimators show that $d \in (0.11, 0.74)$. To balance the tradeoff between bias and variance, estimates for bandwidths $m = T^\alpha$ with $\alpha = 0.525, 0.575$ for GPH and $m = [\frac{T}{8}], [\frac{T}{32}]$ for GSP are also investigated. Results not included here show that estimates of d for both estimators fall within $(-0.05 < d < 0.75)$, the range of consistency for fractional integrated process with long range dependence, Kim and Phillips (2006).

4.1   Model Estimation

Since the returns show serial autocorrelation, we fit an autoregressive structure to the return series:

$$\Phi(L) = (r_t - \mu) = \varepsilon_t \tag{30}$$

where $\Phi(L) = \phi_1 L - \cdots - \phi_p L^p$ as an autoregressive lag polynomial of order p. We assume that the conditional mean

$$\mu_t = \mu + \sum_{i=1}^{p} \phi_i (r_{t-i} - \mu) \tag{31}$$

	Stock index		Stocks	
	FTSE	Bayer	Siemens	VW
std. dev.	1.15	2.31	2.45	2.11
min.	-5.58	-19.42	-10.95	-9.65
max.	5.90	32.99	10.59	10.69
skew.	-0.16	1.10	0.07	-0.01
exc. kurt	2.96	27.71	1.79	2.38
$Q^2(24)$	2329.98 [0.0]	136.64 [0.0]	1724.63 [0.0]	794.20 [0.0]

Table 1: *Descriptive statistics for the daily log-returns for the FTSE stock index* (04.01.2000 − 30.10.2006) *and DAX stocks; Bayer, Siemens and Volkswagen (VW) from* (04.01.2000 − 30.10.2006). $Q^2(24)$ *is the Box-Pierce statistic for remaining serial correlation in the squared standardized residuals using 24 lags, with p-values in brackets.*
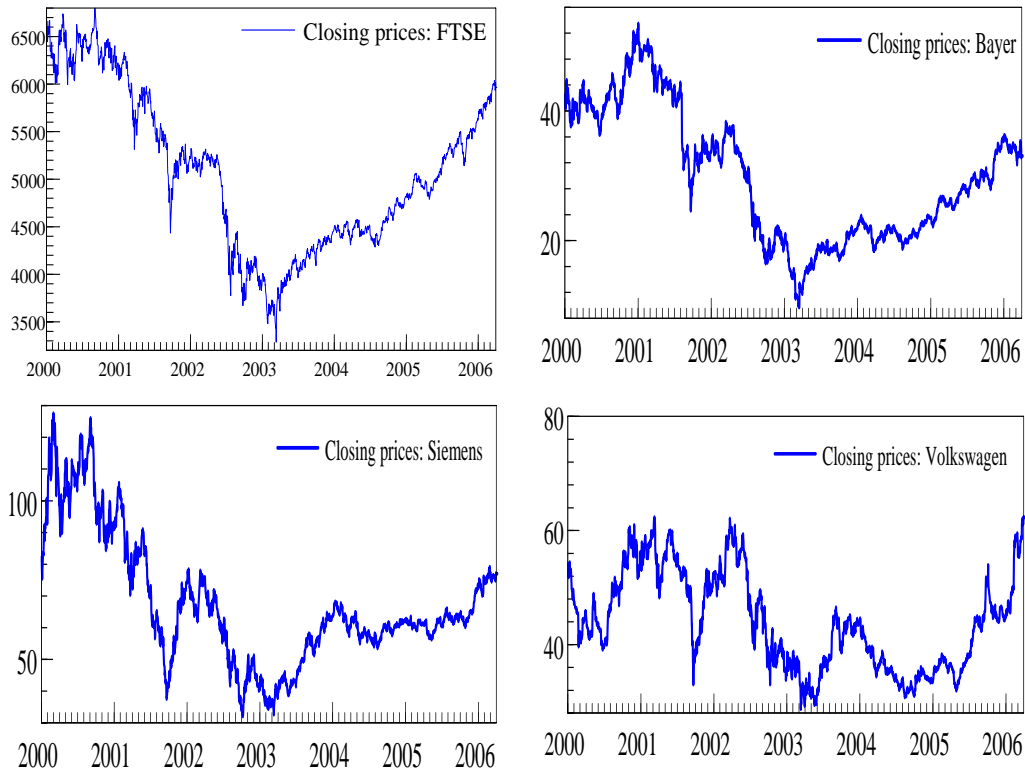
Figure 2: *Daily closing prices for the FTSE index level and DAX stocks; Bayer, Siemens and Volkswagen.*
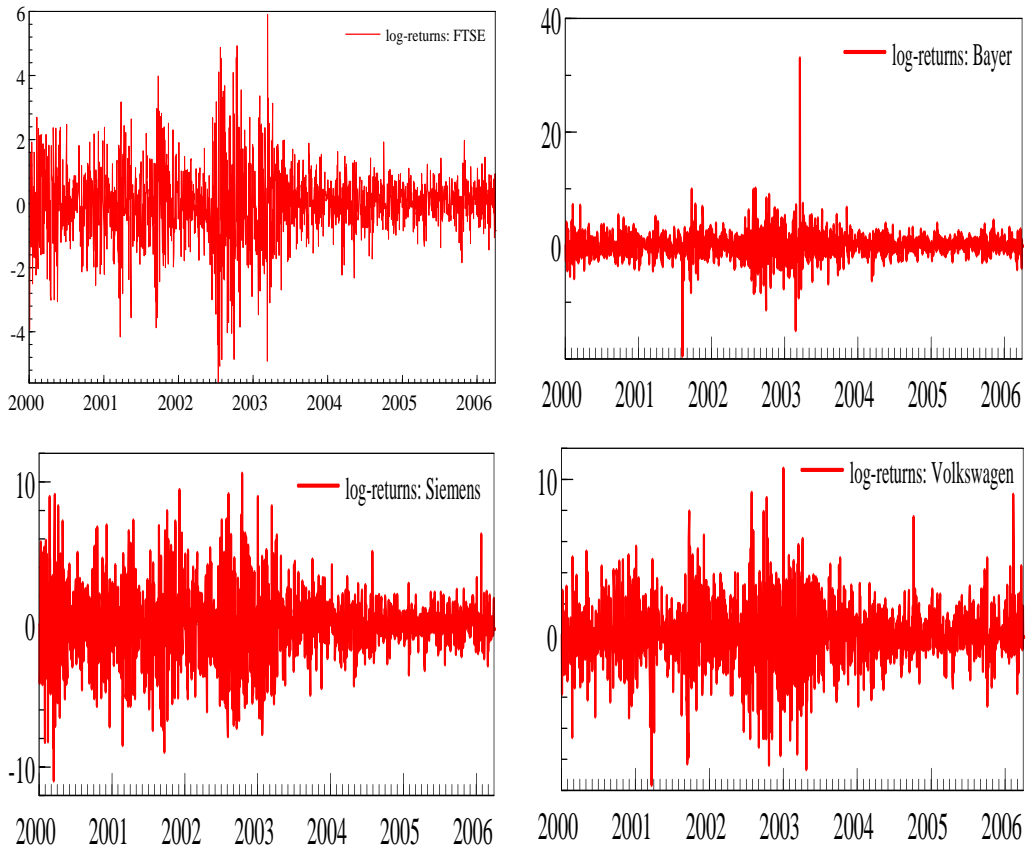
Figure 3: *Daily log-returns for the FTSE index level and DAX stocks; Bayer, Siemens and Volkswagen.*

V/S	$|r_t|$				r_t^2			
m	FTSE	Bayer	Siemens	VW	FTSE	Bayer	Siemens	VW
5	2.41	1.90	3.27	1.86	1.52	0.66	5.40	1.42
10	1.33	1.31	2.09	1.25	0.97	0.56	4.43	0.99
32	0.56	0.58	0.83	0.58	0.43	0.35	2.86	0.50
40	0.47	0.48	0.69	0.49	0.37	0.32	2.62	0.44
110	0.21	0.23	0.31	0.25	0.21	0.20	1.83	0.23

Table 2: *Rescaled Variance test on log-returns for absolute return, $|r_t|$, square returns of the FTSE index and DAX stocks: Bayer, Siemens, Volkswagen (VW). Sample size $T = 1781$ and 5% significance level.*

20
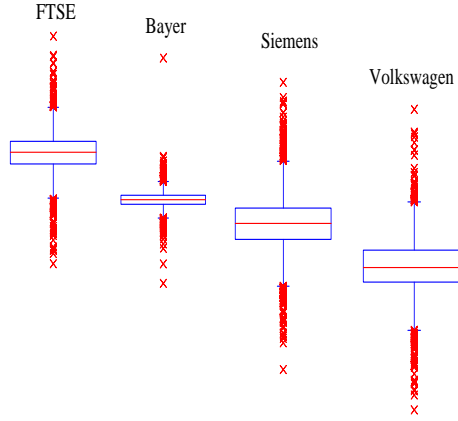
Figure 4: *Boxplots of log-returns of the FTSE index and the DAX stocks, Bayer, Siemens and Volkswagen.*

			$|r_t|$				r_t^2	
	FTSE	Bayer	Siemens	VW	FTSE	Bayer	Siemens	VW
GPH								
$m = T^{0.5}$	0.66	0.67	0.71	0.58	0.49	0.26	0.64	0.52
$m = T^{0.6}$	0.58	0.54	0.60	0.48	0.57	0.35	0.55	0.40
$m = T^{0.8}$	0.35	0.22	0.25	0.25	0.37	0.12	0.24	0.24
GSP								
$m = [\frac{T}{4}]$	0.67	0.24	0.30	0.25	0.57	0.11	0.28	0.24
$m = [\frac{T}{16}]$	0.45	0.44	0.53	0.46	0.44	0.25	0.51	0.39
$m = [\frac{T}{64}]$	0.33	0.65	0.74	0.66	0.35	0.35	0.75	0.70

Table 3: *Log-periodogram regression (GPH), and the Gaussian semiparametric (GSP) estimates for absolute return, $|r_t|$, square returns of the FTSE index and DAX stocks: Bayer, Siemens, Volkswagen (VW). Sample size $T = 1781$.*

follows an $AR(2)$ process in order to capture non-synchronous trading effects. The process specification is expressed as

$$r_t = c + a_1 r_{t-1} + a_2 r_{t-2} + \varepsilon_t \tag{32}$$

$$\varepsilon_t = z_t \sigma_t \tag{33}$$

where z_t is identically and independently distributed with zero mean and unit variance, $\varepsilon_t \sim (0, \sigma_t^2)$. For the conditional variance of ε_t, we fit the $FIAPARCH$ and $HYGARCH$ models under the student-t and skewed student-t distributed innovations. A lag order, $p = q = 1$ is assumed for both models specifications.

Estimation results for the $FIAPARCH$ and $HYGARCH$ models are presented in Tables 4 and 5 respectively. Under skewed student-t distribution, returns of the $FTSE$ equity and Siemens stock are skewed to the left, while Bayer and Volkswagen exhibit right skewness. This result also reveal evidence of the fat-tail phenomenon, given that the tail dependence parameter (ν) is significantly different from zero. Since estimates of δ for $FTSE$, Bayer and Siemens are not significantly different from 1 but significantly different from 2, it is more relevant to model the conditional standard deviation instead of the conditional variance. For Volkswagen stock, modelling the conditional variance is relevant as the δ estimates are significantly different from 1. In addition, γ is positive and significant, indicating leverage effect for negative returns in the conditional variance specification.

For the $HYGARCH$ model results, Table 5, the hyperbolic parameter, $\log(\hat{\alpha}) < 0$ are not significantly different from zero, indicating that the $GARCH$ components are covariance stationary. The AIC and log-likelihood values show that the $FIAPARCH$ model with skewed student-t distributed innovations perform better for all the returns series. This is in line with Degiannakis (2004), that the $FIAPARCH$ models with skewed student-t distributed innovations effectively captures the characteristics of the empirical distribution.

Diagnostic results from the standardized and squared standardized residuals (Q-statistics) and the residual based test (RBD, Tse (2002) for con-

ditional heteroscedasticity do not reject the null hypothesis of a correct model specification for both models. The $FIAPARCH$ model is observed to perform better in modeling the returns than the symmetric ($HYGARCH$) model. It has the advantage in that, it features asymmetry in relationship between the conditional variance and the lagged squared innovations, whatever the chosen error term. Our reference to symmetry and asymmetry is in relation to the distribution of the innovations. In this case, the $FIAPARCH$ model is more adequate to model the negative leverage effect with skewed student-t distributed innovations.

Several criteria have been used to assess the forecasting performance of the models; the Mincer-Zarnowitz (1969) regression, Mean Square Error (MSE), Mean Absolute Prediction Error ($MAPE$) and the logarithmic Loss Function (LL). In-sample one-day-ahead forecast performance of the $HYGARCH$ and $FIAPARCH$ models under the skewed student-t distributed innovations are given in Table 6. Our result shows that the forecast accuracy for the $FIAPARCH$ is marginally better compared to the $HYGARCH$ model.

Figure 5 presents plots of the time path of conditional standard deviation for each log-return series based on the $FIAPARCH$ and $HYGARCH$ models. As observed, the $FIAPARCH$ model is slightly better in estimating the magnitude of volatility than the $HYGARCH$. Our results suggest that such models hold the promise of improved long-run volatility forecast and more accurate pricing of long-term contracts. The implication can therefore be extended to the context of risk management. In what follows, we employ the $FIAPARCH$ model (the $EWMA$ model as benchmark) with skewed student-t distributed innovations for Value-at-Risk and Expected Shortfall.

4.2   Value-at-Risk and Expected Shortfall Predictions

For an overview, we present in Figure 6 the VaR predictions for the $FTSE$ index and DAX stocks returns (Bayer, Siemens, Volkswagen) based on the classical EMA and EDF estimation methods. As observed, these methods do not give satisfactory results since the plots show many returns exceeding the VaR. Though EMA gives better calibration results than EDF, it

	FTSE		Bayer		Siemens		Volkswagen	
	t	Skt	t	Skt	t	Skt	t	Skt
μ	0.01	0.01	0.01	0.02	0.01	0.01	0.02	0.04
	(0.32)	(0.12)	(0.36)	(0.56)	(0.35)	(0.35)	(0.38)	(1.07)
a_1	-0.06	-0.05	0.01	0.01	0.01	0.01	0.05	0.05
	(-2.66)	(-1.96)	(0.56)	(0.57)	(0.63)	(0.67)	(2.18)	(2.28)
a_2	-0.03	-0.03	0.00	0.00	0.02	0.02	-0.02	-0.02
	(-1.21)	(-1.42)	(0.04)	(0.05)	(0.64)	(0.64)	(-1.03)	(-0.97)
ω	1.65	1.12	4.25	4.28	12.03	12.01	5.87	5.97
	(5.22)	(2.94)	(3.94)	(3.94)	(2.50)	(2.50)	(3.10)	(3.09)
d	0.42	0.38	0.41	0.41	0.61	0.61	0.46	0.47
	(8.69)	(9.10)	(6.87)	(6.96)	(7.02)	(6.98)	(7.67)	(7.88)
α_1	0.26	0.28	0.35	0.35	0.20	0.20	0.30	0.29
	(5.88)	(7.48)	(4.98)	(4.78)	(4.11)	(4.11)	(4.68)	(4.60)
β_1	0.60	0.60	0.69	0.69	0.78	0.78	0.65	0.66
	(12.48)	(12.49)	(9.32)	(9.03)	(13.23)	(13.15)	(10.25)	(10.69)
γ	0.93	0.95	0.79	0.78	0.45	0.45	0.21	0.22
	(4.74)	(6.44)	(3.33)	(3.32)	(3.19)	(3.19)	(2.39)	(2.46)
δ	1.16	1.22	1.27	1.27	1.38	1.38	1.80	1.79
	(11.61)	(12.78)	(10.15)	(10.17)	(7.46)	(7.44)	(11.27)	(11.45)
ν	119.20	132.89	6.70	6.67	12.46	12.45	7.12	6.95
	(38.03)	(47.45)	(5.18)	(5.18)	(3.60)	(3.60)	(6.06)	(6.28)
ξ		-0.17		0.02		-0.01		0.07
		(-4.66)		(0.68)		(-0.07)		(2.47)
$Ln(\ell)$	-2292.49	-2291.27	-3580.33	-3580.12	-3750.44	-3750.44	-3631.49	-3628.60
AIC	2.67	2.66	4.03	4.03	4.22	2.77	4.09	4.08
$Q(24)$	14.36	18.49	19.71	19.78	24.33	24.33	31.05	30.87
	(0.70)	(0.73)	(0.60)	(0.59	(0.33)	(0.32)	(0.09)	(0.09)
$Q^2(24)$	19.87	23.21	2.98	3.00	20.96	20.95	18.68	18.77
	(0.33)	(0.38)	(0.99)	(0.99)	(0.52)	(0.52)	(0.66)	(0.65)
$RBD(24)$	22.59	21.20	2.30	2.32	21.07	21.08	17.80	17.93
	(0.54)	(0.62)	(1.00)	(1.00)	(0.63)	(0.63)	(0.81)	(0.80)

Table 4: *Estimation results from* $AR(2) - FIAPARCH(1,d,1)$ *model using student-t, and skewed student-t distribution. t-statistics of corresponding tests in brackets.* $Ln(\ell)$ *is the value of the maximized log-likelihood.* $Q(24)$ *and* $Q^2(24)$ *are the Box-Pierce statistic for remaining serial correlation in the standardized and squared standardized residuals respectively.* $RBD(24)$ *is the residual based diagnostic for conditional heteroscedasticity, using* 24 *lags with p-values in parenthesis.*

	FTSE		Bayer		Siemens		Volkswagen	
	t	Skt	t	Skt	t	Skt	t	Skt
μ	0.04	0.03	0.05	0.06	0.05	0.06	0.03	0.06
	(2.29)	(2.07)	(1.30)	(1.54)	(1.42)	(1.46)	(0.67)	(1.41)
a_1	-0.08	-0.09	0.01	0.02	0.02	0.02	0.05	0.05
	(-3.12)	(-3.79)	(0.56)	(0.63)	(0.67)	(0.71)	(2.13)	(2.28)
a_2	-0.04	-0.05	-0.01	-0.01	0.01	0.01	-0.02	-0.02
	(-1.64)	(-2.03)	(-0.27)	(-0.35)	(0.41)	(0.42)	(-1.02)	(-0.91)
ω	0.02	0.02	0.01	0.02	0.01	0.01	0.06	0.05
	(1.69)	(1.66)	(0.70)	(69)	(0.20)	(0.17)	(0.73)	(0.49)
d	0.60	0.60	1.13	1.14	0.65	0.65	0.48	0.46
	(7.01)	(6.49)	(7.34)	(7.55)	(3.40)	(3.55)	(3.84)	(3.73)
α_1	0.06	0.08	-0.09	-0.10	0.16	0.16	0.29	0.29
	(1.20)	(1.42)	(-0.83)	(-0.91)	(1.77)	(1.83)	(3.63)	(3.57)
β_1	0.60	0.60	0.97	0.97	0.79	0.79	0.66	0.65
	(7.32)	(6.48)	(23.89)	(24.88)	(7.23)	(7.59)	(7.73)	(7.63)
ξ		-0.16		0.03		0.02		0.07
		(-4.45)		(0.97)		(0.51)		(2.36)
ν	33.97	58.97	5.84	5.80	11.13	11.12	6.76	6.56
	(1.25)	(5.22)	(6.09)	(6.10)	(3.98)	(2.79)	(6.14)	(6.34)
$\log(\hat{\alpha})$	-0.02	-0.03	-0.01	-0.01	0.00	0.00	0.01	0.02
	(-0.90)	(-1.12)	(-0.72)	(-0.71)	(0.26)	(0.30)	(0.15)	(0.36)
$Ln(\ell)$	-2319.74	-2310.46	-3599.01	-3598.57	-3765.05	-3764.93	-3636.06	-3633.42
AIC	2.70	2.67	4.05	4.05	4.24	4.24	4.05	4.02
$Q(24)$	17.65	19.02	20.32	20.37	22.31	22.31	29.70	29.39
	(0.72)	(0.64)	(0.56)	(0.55)	(0.44)	(0.44)	(0.13)	(0.13)
$Q^2(24)$	18.37	18.25	2.42	2.41	22.06	22.18	18.14	18.54
	(0.68)	(0.64)	(0.99)	(0.99)	(0.45)	(0.44)	(0.69)	(0.67)
$RBD(24)$	17.76	17.65	2.24	2.23	22.16	22.31	16.96	17.59
	(0.81)	(0.82)	(1.00)	(1.00)	(0.56)	(0.56)	(0.85)	(0.82)

Table 5: *Estimation results from $AR(2) - HYGARCH(1, d, 1)$ model using student-t, and skewed student-t distribution. t-statistics of corresponding tests in brackets. $Ln(\ell)$ is the value of the maximized log-likelihood. $Q(24)$ and $Q^2(24)$ are the Box-Pierce statistic for remaining serial correlation in the standardized and squared standardized residuals respectively. $RBD(24)$ is the residual based diagnostic for conditional heteroscedasticity, using 24 lags with p-values in parenthesis.*
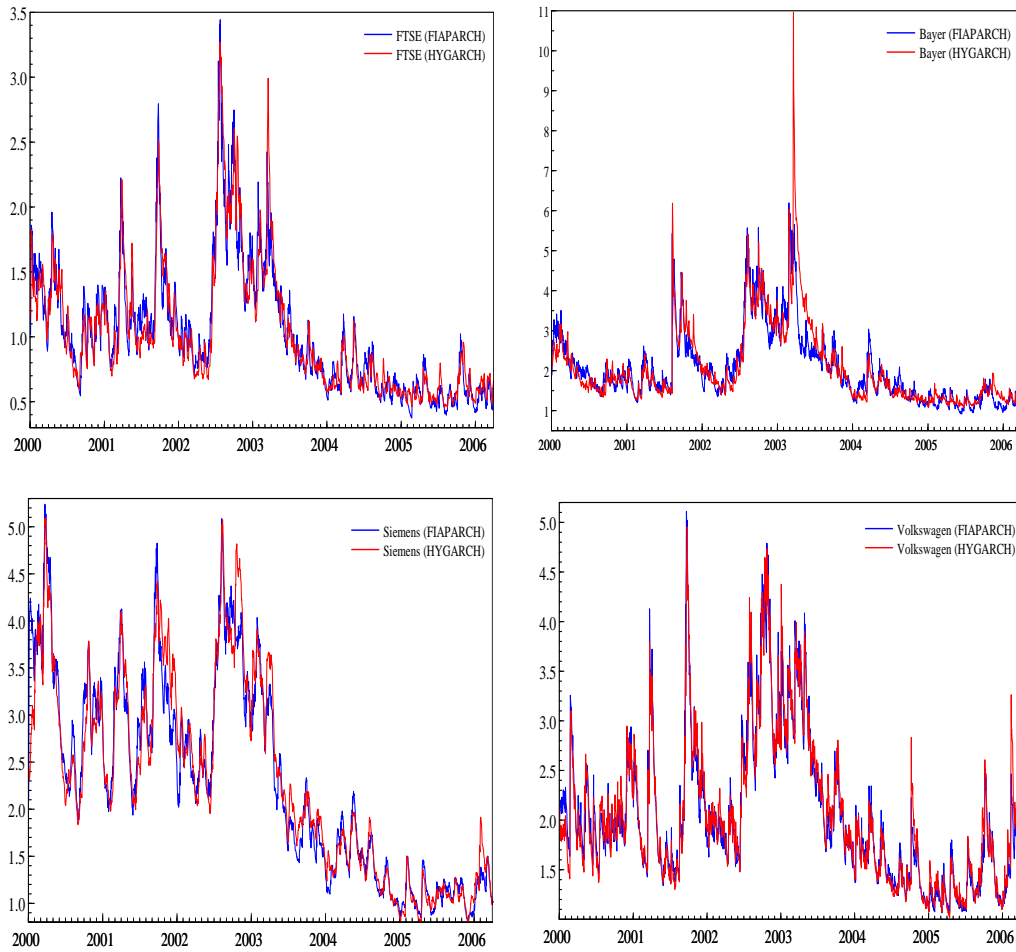
Figure 5: *Plots of the time path of conditional standard deviation for each log-return series based on the FIAPARCH and HYGARCH models under the skewed student-t distributed errors.*

HYGARCH	c	β	R^2	MSE	MAPE	LL
FTSE	0.02	0.99	0.15	1.24	15300	7.22
Bayer	0.36	0.67	0.03	18.66	2479	7.62
Siemens	1.14	0.42	0.01	18.47	103.40	6.99
VW	2.20	0.22	0.002	59.37	145.50	7.69
FIAPARCH	c	β	R^2	MSE	MAPE	LL
FTSE	0.04	1.13	0.18	1.21	11600	7.22
Bayer	-0.94	1.50	0.11	17.02	1913	6.79
Siemens	0.50	0.80	0.03	17.76	95.47	6.69
VW	1.94	0.35	0.003	58.19	131.50	7.30

Table 6: *In sample one-step-ahead forecasting performance of the HYGARCH and FIAPARCH models on log returns. Forecast criteria are the Mincer-Zarnowitz regression (c, β are regression coefficients, R^2, the determination coefficient of regression for the model), MSE is the Mean Absolute Prediction Error, MAPE is the Root Mean Square Error and LL, the logarithmic Loss Function.*

cannot effectively be applied to model long range dependence because of its exponentially decreasing memory.

We apply the $FIAPARCH$ model under the skewed student-t distributed innovations to calculate the Value-at-Risk and Expected Shortfall. Backtesting VaR and ESF results for the $FIAPARCH$ models are presented in Tables 7 to 10. The computed values are the failure rate, Kupiec LR test and P-values, the Expected Shortfall (ESF) with significance level, $\alpha = 0.01, 0.05$ and $\alpha = 0.95, 0.99$ for long and short positions respectively. The failure rate for short position refers to the percentage of positive returns larger than the VaR prediction, while the failure rate for the long position is the percentage of negative returns smaller than the VAR prediction. Results show that the VaR for the $FIAPARCH$ model performs very well for the one-day as well as for the five-days time horizon. With more consecutive choice of α, the model captures the fat tails exhibited in the returns very well.

The one-step ahead VaR prediction plots, Figure 7 for the $FIAPARCH$ model shows that the model does well in predicting critical loss for long and short positions. The test of VaR is not rejected for all levels. For

FTSE	$h = 1$ day		$h = 5$ days	
Short position				
α Quantile	0.950	0.990	0.950	0.990
Failure rate $(\frac{N}{T})$	0.948	0.990	0.948	0.991
Kupiec-LR_1	0.039	0.092	0.048	0.297
P-value	0.842	0.761	0.825	0.587
ESF	1.915	2.409	1.915	2.511
Long position				
α Quantile	0.050	0.010	0.050	0.010
Failure rate $(\frac{N}{T})$	0.055	0.008	0.054	0.008
Kupiec-LR_1	0.916	0.307	0.761	0.297
P-value	0.338	0.579	0.382	0.585
ESF	-2.241	-2.587	-2.257	-2.587

Table 7: *In-sample Value-at-Risk (VaR) and Expected Shortfall (ESF) evaluation under the skewed student-t distribution - FIAPARCH for FTSE log-returns.*

example, the proportion of exceptions at 1% is about 0.8% for $FTSE$, $Bayer$, $Siemens$ and 1.1% for Volkswagen. Our results indicate that, models that account for asymmetries as well as fractional integrated parametrization of the volatility process, perform better in predicting VaR and ESF for short and long positions than short memory models. This suggest that for proper risk valuation, the degree of persistence of the process should be investigated and appropriate model that incorporate the existence of such characteristics be taken into account.

Predicted five-days-ahead out-of-sample VaR and ESF forecast results for the $EWMA$ and $FIAPARCH$ models are presented in Tables 11 and 12 respectively. These forecasts are based on a 'sliding window', updating the model parameters every 50 observations. Though not as good as the in-sample forecasts, the results show that the $FIAPARCH$ model specification performs well for out-of-sample VaR predictions.
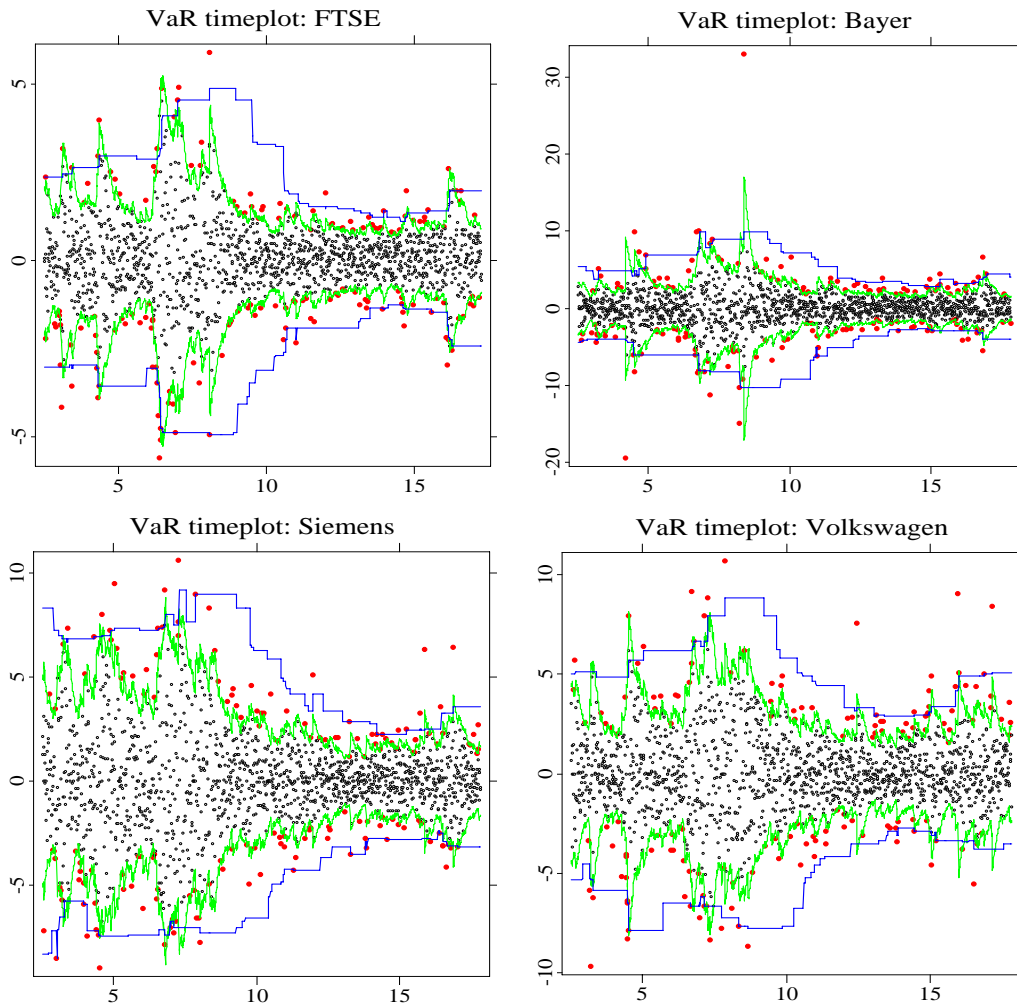
Figure 6: *Time plots of VaR forecasts at* 99% *and* 1% *level using the EDF (blue) and EMA (green) models (with decay factor,* $\lambda = 0.94$ *and a rolling window of* 250 *days) for the returns series. Returns which exceed the VaR are marked in red.*
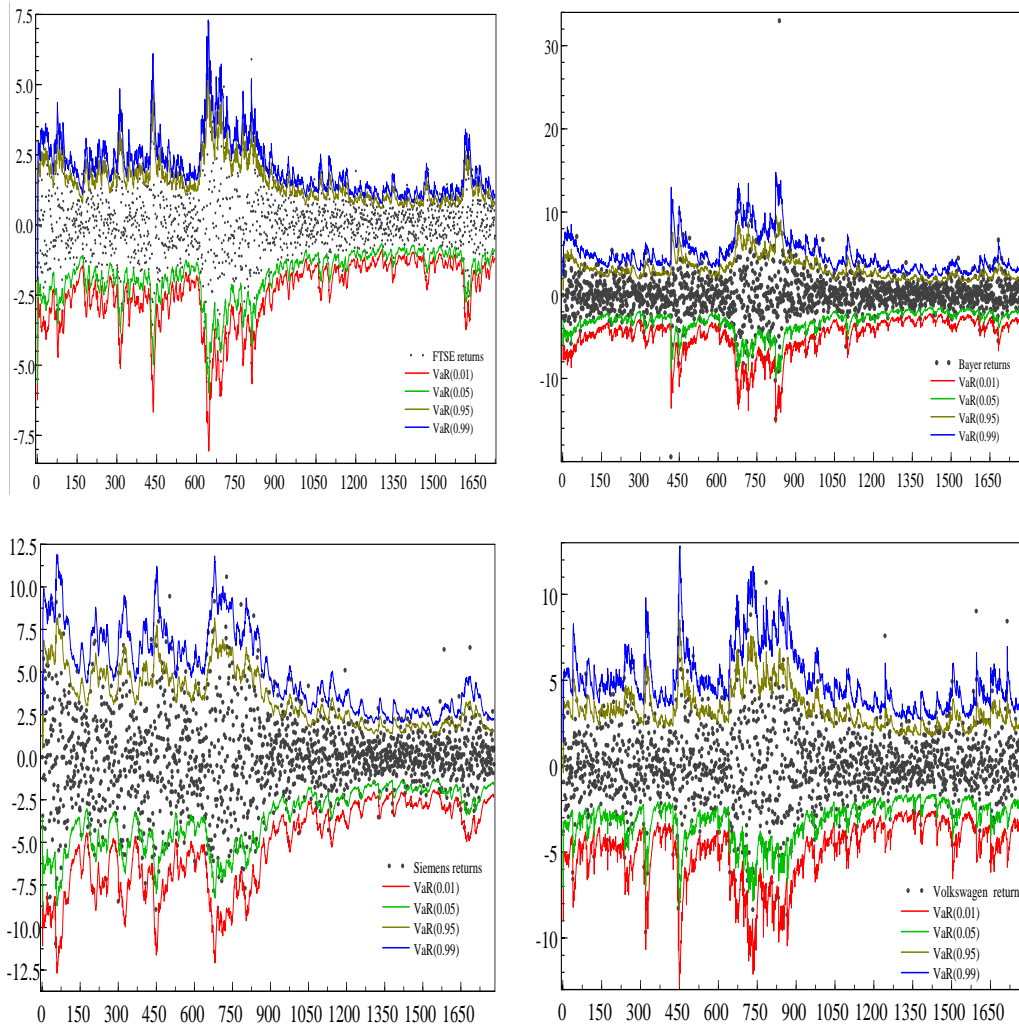
Figure 7: *One-day-ahead VaR forecast plots for the $FTSE$ and DAX stocks log-returns (Bayer, Siemens, Volkswagen) with $FIAPARCH$ model under skewed student-t distributed innovations. The α quantiles for long and short positions are $\alpha = 0.01, 0.05$ and $\alpha = 0.95, 0.99$ respectively.*

Bayer	$h = 1$ day		$h = 5$ days	
Short position				
α Quantile	0.950	0.990	0.950	0.990
Failure rate ($\frac{N}{T}$)	0.951	0.992	0.952	0.992
Kupiec-LR_1	0.107	0.884	0.093	0.867
P-value	0.742	0.347	0.759	0.351
ESF	4.556	7.916	4.556	7.916
Long position				
α Quantile	0.050	0.010	0.050	0.010
Failure rate ($\frac{N}{T}$)	0.052	0.008	0.051	0.008
Kupiec-LR_1	0.186	0.470	0.056	0.457
P-value	0.665	0.492	0.811	0.498
ESF	-4.531	-7.131	-4.526	-7.131

Table 8: *In-sample Value-at-Risk (VaR) and Expected Shortfall (ESF) evaluation under the skewed student-t distribution - $FIAPARCH$ for Bayer log-returns. * is the 5% confidence level.*

Siemens	$h = 1$ day		$h = 5$ days	
Short position				
α Quantile	0.950	0.990	0.950	0.990
Failure rate ($\frac{N}{T}$)	0.952	0.992	0.953	0.992
Kupiec-LR_1	0.301	0.884	0.407	0.867
P-value	0.583	0.347	0.523	0.351
ESF	4.737	5.781	4.775	5.781
Long position				
α Quantile	0.050	0.010	0.050	0.010
Failure rate ($\frac{N}{T}$)	0.053	0.008	0.052	0.009
Kupiec-LR_1	0.417	0.190	0.314	0.182
P-value	0.518	0.662	0.574	0.669
ESF	-4.468	-5.075	-4.492	-5.075

Table 9: *In-sample Value-at-Risk (VaR) and Expected Shortfall (ESF) evaluation under the skewed student-t distribution - $FIAPARCH$ for Siemens log-returns.*

Volkswagen	$h = 1$ day		$h = 5$ days	
Short position				
α Quantile	0.950	0.990	0.950	0.990
Failure rate $(\frac{N}{T})$	0.951	0.992	0.951	0.992
Kupiec-LR_1	0.107	0.884	0.093	0.457
P-value	0.742	0.347	0.759	0.498
ESF	4.087	6.166	4.093	6.050
Long position				
α Quantile	0.050	0.010	0.050	0.010
Failure rate $(\frac{N}{T})$	0.052	0.011	0.052	0.012
Kupiec-LR_1	0.186	0.549	0.206	0.564
P-value	0.665	0.458	0.649	0.452
ESF	-4.359	-5.738	-4.359	-5.738

Table 10: *In-sample Value-at-Risk (VaR) and Expected Shortfall (ESF) evaluation under the skewed student-t distribution - FIAPARCH for Volkswagen log-returns.*

$h = 5$ days	**FTSE**		**Bayer**		**Siemens**		**VW**	
Short position	0.95	0.99	0.95	0.99	0.95	0.99	0.95	0.99
Failure rate $(\frac{N}{T})$	0.95	0.99	0.94	0.99	0.95	0.97	0.93	0.96
Kupiec-LR_1	0.16	0.09	0.50	0.80	0.01	2.01	1.65	10.39
P-value	0.68	0.75	0.47	0.36	0.91	0.15	0.19	**0.00***
ESF	1.59	1.71	3.45	4.32	3.00	4.46	4.12	5.14
Long position	0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
Failure rate $(\frac{N}{T})$	0.06	0.02	0.04	0.02	0.04	0.01	0.02	0.00
Kupiec-LR_1	1.04	3.64	0.01	2.01	0.50	0.10	7.97	1.14
P-value	0.30	**0.05***	0.91	0.15	0.47	0.74	**0.00***	0.28
ESF	-1.68	-1.80	-2.86	-3.88	-2.89	-3.23	-3.73	-5.54

Table 11: *EWMA: backtesting results for five-day-ahead out-of-sample VaR and ESF forecasts for the FTSE and DAX stocks log-returns. * indicates a rejection of the model used.*

$h = 5$ days	FTSE		Bayer		Siemens		VW	
Short position	0.95	0.99	0.95	0.99	0.95	0.99	0.95	0.99
Failure rate $\left(\frac{N}{T}\right)$	0.95	0.98	0.93	0.99	0.93	0.97	0.92	0.97
Kupiec-LR_1	0.01	0.10	0.56	0.09	0.50	2.00	3.26	5.60
P-value	0.91	0.74	0.45	0.75	0.45	0.15	0.07	**0.01***
ESF	1.28	1.51	3.25	5.03	2.77	4.46	3.97	5.43
Long position	0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
Failure rate $\left(\frac{N}{T}\right)$	0.05	0.01	0.04	0.01	0.05	0.01	0.02	0.00
Kupiec-LR_1	0.03	0.10	0.50	0.10	0.03	0.10	2.87	1.14
P-value	0.85	0.74	0.47	0.74	0.85	0.74	0.09	0.28
ESF	-1.21	-1.24	-3.14	-4.54	-2.82	-3.37	-3.19	-5.54

Table 12: *FIAPARCH: backtesting results for five-days-ahead out-of-sample VaR and ESF forecasts for the FTSE and DAX stocks log-returns.* * *indicates a rejection of the model used.*

4.3 Summary

In this paper, we applied Value-at-Risk and Expected Shortfall to valuate financial market risk when there is long range dependence. We examined the British equity portfolio, $FTSE100$ and three stocks of the German DAX index (Bayer, Siemens and Volkswagen). Long range dependence test and estimation is carried out using the rescaled variance test of Geweke and Porter-Hudak (1983) (GPH) and two estimators, the log-periodogram regression (GPH), Geweke and Porter-Hudak (1983) and the Gaussian semi-parametric (GSP), Robinson (1995a) estimators. With long memory in the volatility process, we fitted the ($HYGARCH$) model of Davidson (2004) that can accommodate distributional characteristics such as volatility clustering, leptokurtosis, long memory and the ($FIAPARCH$) of Tse (1998), that can well accommodate volatility clustering, leptokurtosis, asymmetry and long memory. In particular, we estimated these models with different error distribution assumptions, the student-t and skewed student-t. Our results show that models that account for asymmetries in the volatility specifications as well as fractional integrated parametrization of the volatility process, perform better in predicting the one-step ahead as well as five-step ahead VaR and ESF for short and long positions than short memory models. Such models hold the promise of improved long-run volatility forecast and more accurate pricing of long-term contracts. We extended this implication to the context of financial risk management. Our results suggest that for proper risk val-

uation of equity or portfolio of stocks, the degree of persistence should be investigated and appropriate modelling that incorporate the existence of such characteristic be taken into account.

References

Aloui, C., Abaoub, E., Bellalah, M. (2005): Long Range Dependence on Tunisian Stock Market Volatility, *International Journal of Business* **10**, no. 4, 1–26.

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

Angelidis, T. and Degiannakis. S. (2006): Backtesting VaR models. An Expected Shortfall Approach, *Athens University of Economics and Business, Department of statistics, technical Report*, 223.

Baillie, R. T., Bollerslev, T., Mikkelson, H. (1996): Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **14**: 3–30.

Beran, J., (1994): Statistics for long memory processes. *Chapman & Hall*, New York.

Billo, M. and Pelizzon, L. (2000): Value-at-Risk: A multivariate switching regime approach. *Journal of Empirical Finance*, **7**, 531–554.

Bollerslev, T. (1986): Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.

Bollerslev, T. (1987): A conditional heteroskedastic times series model for speculative prices and rate of return. *Review of Economics and Statistics*, **69**, 542–547.

Bollerslev, T. and Mikkelsen, H. (1996): Modelling and Pricing Long Memory in Stock Market Volatility. *Journal of Econometrics*, **73**, 151–184.

Brooks, C. and Persand, G. (2003a): The effect of asymmetries on stock index return VaR-at-Risk estimates. *Journal of Risk Finance*, Winter, 29–42.

Brooks, C., Burke, S. and Persand, G. (2005): Autoregressive Conditional Kurtosis, *Journal of Financial Econometrics*, **3**, 399-421

Chung, C. F. (1999): Estimating the Fractionally Integrated GARCH Model. *working paper*, National Taïwan Univerity.

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

Davidson, J. (2004): Moments and memory properties of linear conditional heteroskedasticity models, and a new model. *Journal of Business and Economic Statistics*, **22**, 16–29.

Degiannakis, S. (2004): Volatility forecasting: Evidence from a fractional integrated asymmetric power ARCH skewed-t model. Applied Financial Economics, **14**, 1333–1342.

Ding, Z. and Granger, C.W.J. (1996): Modelling Volatility Persistence of Speculative Returns. A New Approach. *Journal of Econometrics*, **73**, 185–215.

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

Doornik, J.A. (2006a): An Introduction to OxMetrics 4. A software System for Data Analysis and Forecasting. *Timberlake Consultant Ltd.*, first edn.

Doornik, J.A., Ooms, M. (2004): Inference and forecasting for ARFIMA models with an Application to US and UK inflation. *Studies in Nonlinear Dynamics and Econometrics 8*, No. **2**, Article 14.

Engel, R.F. and Bollerslev, T. (1986): Modeling the Persistence of Conditional Variances. *Econometric Reviews*, **5**, 1–5.

Giot, P. and Laurent, S. (2004): Modelling Daily Value-at-Risk Using Realized Volatility and ARCH Type Models. *Journal of Empirical Finance*, **11**, 379-398.

Guermat, C. and Harris, R. D. F. (2002): Robust conditional variance estimation and value-at-risk. *Journal of Risk*, Vol. **4**, No.2

Geweke, Porter-Hudak (1983): The Estimation and Application of Long-Memory Time Series Models. *Journal of Time Series Analysis*, **4**, 221-238.

Giraitis, L., Kokoszka, P. and Leipus, R. (2000): Stationary ARCH models: dependence structure and Central Limit Theorem. *Econometric Theory*, **16**, 3-22.

Giraitis, L., Kokoszka, P. and Leipus, R. (2001): Testing for long memory in the presence of a general trend. *Journal of Applied Probability*, **38**, 1033–1054.

Giraitis, L., Kokoszka, P., Leipus, R. and Teyssière, G. (2003): Rescaled Variance and Related Tests for Long Memory in Volatility and Levels. *Journal of Econometrics*, **112**, 265–294.

Giraitis, L., Kokoszka, P., Leipus, R. and Teyssière, G. (2000): Semiparametric Estimation of the Intensity of Long-Memory in Conditional Heteroskedasticity. *Statistical Inference for Stochastic Processes*, **3**, 113-128.

Granger, C.W.J. (1980): Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, **14**: 261–279.

Granger, C.W.J. and Ding, Z. (1995): Some properties of absolute returns: An alternative measure of risk. *Annales d'Économie et de Statistique*, **40**, 67–91.

Granger, C.W.J. and Ding, Z. (1996): Varieties of Long-Memory Models. *Journal of Econometrics*, **73**, 61–77.

Hassler, U. (1993a): Regression of spectral estimators with fractionally integrated time series. *Journal of Time Series Analysis*, **14**, 339–349.

Hendricks, D. (1996): Evaluation of Value-at-risk Models using Historical Data, *Federal Reserve bank of New York Economic Policy Review*, April 1996.

Henry, O. T., (2002): Long Memory in Stock Returns: Some International Evidence, *Applied Financial Economics*, **12**, 725–729.

Herzberg, M. and Sibbertsen, P. (2004): Pricing of options under different volatility models. *Technical Report/ Universität Dortmund, SFB 475, Komplexitätsreduktion in Multivariaten Datenstrukturen*, 2004, **62**.

Hurvich, C., Doe, R. and Brodsky, J. (1998): The Mean Square Error of Geweke and Porter-Hudak's estimator of the Memory Parameter of a Long-Memory time series. *Journal of Time Series Analysis*, **19**, 19–46.

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

Kim, C.S., and Phillips, P.C.B. (2006): Log periodogram regression in the nonstationary case. *Cowless foundation Discussion paper, no. 1587*, Yale University.

Kupiec, P. (1995): Techniques for Verifying the Accuracy of Risk Management Models, *Journal of Derivatives*, **2**, 173–184 .

Lambert, P. and Laurent, S. (2001): Modeling financial times series using garch-type models and a skewed student density. *Université de Liége.*

Laurent, S. and Peters, J. P., (2002): Garch 2.2: An Ox Package for estimating and forecasting Various ARCH Models. *Journal of Economic Surveys*, **16**(3), 447–485.

Mincer, J. and Zarnowitz, V. (1969): The Evaluation of Economic Forecasts. Economic Forecasts and Expectations in J. Mincer *National Bureau of Economic Research*, New York.

Robinson, P.M. (1990): Time series with strong dependence. *Advances in Econometrics, 6th World Congress, Cambridge*, Cambridge University Press.

Robinson, P.M. (1995a): Gaussian Semiparametric Estimation of Long-Range Dependence. *Annals of Statistics*, **23**, 1630–1661.

Lobato, I. and P.M. Robinson (1998): A Nonparametric Test for I(0). *Review of Economic Studies*, **65**(3), 475-495.

Phillips, P.C.B. and Shimotsu, K. (2004): Local Whittle estimation in non-stationary and unit root cases. *Annals of Statistics*,**32**, 656-692.

Robinson, P.M., Henry, M., (1999): Long and Short Memory Conditional Heteroskedasticity in Estimating the Memory Parameter in Levels. *Economic Theory*, **15**, 299-336.

Scaillet, O. (2004): Nonparametric Estimation and Sensitivity Analysis of Expected Shortfall, *Mathematical Finance*, **14**, 115–129.

So, M.k.P. (2000): Long-term memory in stock market volatility. *Applied Financial Economics*, **10**, 519–524.

So, M.k.P. and Yu, P.L.H. (2006): Empirical analysis of GARCH models in Value at Risk estimation. *Journal of International Markets, Institution and Money*, **16(2)**, 180–197

Tse, Y.K (1998): The Conditional Heteroskedasticity of the Yen-Dollar Exchange Rate. *Journal of Applied Econometrics*, **193**, 49–55

Tse, Y.K (2002): Residual-based Diagnostics for Conditional Heteroskedasticity Models. *Econometrics Journal*, **5**, 358–373

SFB 649 Discussion Paper Series 2008

For a complete list of Discussion Papers published by the SFB 649,
please visit http://sfb649.wiwi.hu-berlin.de.

001   "Testing Monotonicity of Pricing Kernels" by Yuri Golubev, Wolfgang
        Härdle and Roman Timonfeev, January 2008.
002   "Adaptive pointwise estimation in time-inhomogeneous time-series
        models" by Pavel Cizek, Wolfgang Härdle and Vladimir Spokoiny,
        January 2008.
003   "The Bayesian Additive Classification Tree Applied to Credit Risk
        Modelling" by Junni L. Zhang and Wolfgang Härdle, January 2008
004   "Independent Component Analysis Via Copula Techniques" by Ray-Bing
        Chen, Meihui Guo, Wolfgang Härdle and Shih-Feng Huang, January
        2008.
005   "The Default Risk of Firms Examined with Smooth Support Vector
        Machines" by Wolfgang Härdle, Yuh-Jye Lee, Dorothea Schäfer
        and Yi-Ren Yeh, January 2008.
006   "Value-at-Risk and Expected Shortfall when there is long range
        dependence" by Wolfgang Härdle and Julius Mungo, Januray 2008.