A Monte Carlo Simulation Approach to Forecasting Multi-period Value-at-Risk and Expected Shortfall Using the FIGARCH-skT Specification

Degiannakis, Stavros and Dent, Pamela and Floros, Christos

Department of Economics, Portsmouth Business School, University of Portsmouth, Postgraduate Department of Business Administration, Hellenic Open University

2014

A Monte Carlo Simulation Approach to Forecasting Multi-period Value-at-Risk and Expected Shortfall Using the FIGARCH-skT Specification

Stavros Degiannakis[∞,*,**], Pamela Dent[*] and Christos Floros[*]

[*]Department of Economics, Portsmouth Business School, University of Portsmouth, Richmond Building, Portland Street, Portsmouth, UK, PO1 3DE

[**] Postgraduate Department of Business Administration, Hellenic Open University, Aristotelous 18, Greece, 26 335,

Abstract

In financial literature, Value-at-Risk (VaR) and Expected Shortfall (ES) modelling is focused on producing 1-step ahead conditional variance forecasts. The present paper provides a methodological contribution to the multi-step VaR and ES forecasting through a new adaptation of the Monte Carlo simulation approach for forecasting multi-period volatility to a fractionally integrated GARCH framework for leptokurtic and asymmetrically distributed portfolio returns. Accounting for long memory within the conditional variance process with skewed Student-t (skT) conditionally distributed innovations, accurate 95% and 99% VaR and ES forecasts are calculated for multi-period time horizons. The results show that the FIGARCH-skT model has a superior multi-period VaR and ES forecasting performance.

Keywords: Expected Shortfall, FIGARCH, Forecasting, stock indices, skewed Student-t, Volatility, Long Memory, Value-at-Risk, VaR.

JEL Classifications: G17; G15; C15; C32; C53.

[∞] Corresponding author, email: stavros.degiannakis@gmail.com.

1. Introduction

Value-at-Risk (VaR) is an important tool in risk measurement and the management of the financial assets. Originally used internally by financial institutions to assess risk, VaR assumed greater significance when the Basel Committee encouraged its use through the 1996 Market Risk Amendment to the 1988 Basel Accord (Basel, 1988; Basel, 1996). Subsequently, the Basel Committee has refined the regulations relating to the use of VaR, allowing greater flexibility for certain financial institutions to use their own internal VaR models subject to the models being approved by the regulator (Basel, 2006).

VaR quantifies the maximum loss for a portfolio of assets under normal market conditions over a given period of time and at a certain confidence level. Although financial institutions have flexibility over the model which is used to estimate VaR, the regulations prescribe that they use up to one year of historical data to calculate the daily VaR for their positions. This daily VaR should be up scaled to a 10-day VaR figure to represent the banks having at least a 10-day holding period for any given position. The recent financial crash has highlighted the importance and need for reliable models to predict VaR, and has led to further amendments to the regulations, which now require financial institutions to additionally calculate a 'stressed value-at-risk' measure using a one year data period in which the bank incurred significant losses (Basel, 2009). Expected Shortfall (ES) is an alternative to VaR that is more sensitive to the shape of the loss distribution in the tail of the distribution. ES quantifies the expected value of the loss, given that a VaR violation has occurred.

Within the literature the ability of a variety of increasingly complex models (both parametric and non-parametric) to estimate and forecast VaR has been tested. These models can account variously for certain features of financial asset returns such as heteroskedasticity, asymmetry or leverage effects, leptokurtic distribution and long memory (hyperbolic decline of the conditional variance); see Alexander (2008) for more details. The various competing models have been compared using a range of distributions for the standardised residuals (normal, Student-*t*, skewed Student-*t*, generalized error distribution, stable Paretian, exponential generalized beta), across a number of markets for different levels of statistical significance, often for both long and short positions; see González-Rivera et al. (2004), Xekalaki and Degiannakis (2010) for more details.

At present, the findings in the literature are highly inconsistent as to which is the optimal model for estimating VaR. The best model appears to vary, amongst other factors, with the length of the data series, the market for which VaR is being estimated and the assumptions regarding the distribution of the standardised residuals (Angelidis et al., 2004).

Furthermore, a model found to be superior for estimating VaR for long positions may not be optimal for estimating VaR for short positions due to the asymmetric distribution of financial returns (Shao et al., 2009).

The empirical success of the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) framework (Engle, 1982 and Bollerslev, 1986) to model high frequency volatility has been widely highlighted, with many papers focussing on the selection of the optimal GARCH specification in order to calculate and predict VaR (see, for example, Giot and Laurent, 2003a, 2003b, 2004, Caporin, 2008, Tang and Shieh, 2006, McMillan and Kambouroudis, 2009). Literature provides evidence that among the simple models, the GARCH(1,1) model is the most adequate one. Thus, our intention is to compare the baseline model with a more complex specification to allow an assessment of the trade-off between complexity and accuracy. Berkowitz and O'Brien (2002), in order to assess the performance of the banks structural models, compare their VaR forecasts with those from a GARCH model of the banks P&L volatility. They provide evidence that the banks structural VaR models do not provide forecasts superior to a simple GARCH model of P&L volatility[1].

Grané and Veiga (2008) demonstrate that long memory models outperform short memory GARCH specifications, but like the majority of VaR studies, limit their backtesting to forecasting horizon of just one trading day. By contrast, financial institutions are required by the Basel Committee to calculate the VaR of their positions for at least a 10-day holding period in order to calculate their minimum capital risk requirements (Basel, 2009). Although the Basel Committee suggest that 10-day VaR may be calculated by augmenting 1-day VaR using the square root of time rule[2], Wang et al. (2011), Engle (2004) and Danielsson (2002) criticise this technique on the basis that it makes the invalid assumption that the returns are independently and identically normally distributed and that volatilities over time are constant. Danielsson and Zigrand (2006) show that the square root of time scaling rule can lead to an underestimation of market risk, especially for longer time horizons. Beltratti and Morana (1999) find for horizons of one, five and ten days, that the FIGARCH model produces similar VaR forecasts to the simpler GARCH model, when the multi-period forecasts have been constructed based on the square-root-of-time rule.

Hartz et al. (2006) employ a re-sampling technique based on the bootstrap and bias correction step to improve the multi-period VaR forecasts produced by the simple normally

---

[1] The GARCH model, for lower VaRs, is better at predicting changes in volatility and permits comparable risk coverage with less regulatory capital.

[2] To account for the non-linear price characteristics of options contracts, financial institutions are expected to move towards calculating a full 10-day VaR for positions involving such contracts.

distributed AR(1)-GARCH model. They employ another standard multi-period forecasting technique of iterating the conditional mean and conditional variance specifications, using expected values where the returns or innovations are inestimable. Brooks and Persand (2003) use a similar technique to investigate methods of evaluating multi-period volatility forecasts produced by a range of GARCH family and other linear models. By contrast, Kinateder and Wagner (2010) show that a scaling-based GARCH-LM technique produces superior VaR forecasts to a benchmark fully parametric GARCH model utilising the Drost-Nijman (1993) formula for multiday volatility forecasts, especially for the five and ten day horizons.

Multi-period VaR may be estimated using a variety of techniques, including parametric or variance-covariance approaches, non-parametric approaches (e.g. historical simulation), semi-parametric approaches (e.g. extreme value theory) and Monte Carlo simulation (Dionne et al., 2009). For example, Semenov (2009) proposes a historical simulation technique which allows the accurate estimation of 1-day and 10-day VaR figures conditional on the historical sensitivity of assets returns (within a portfolio) to various macroeconomic factors (risk factor betas) over a period of time. Dionne et al. (2009) use a Monte Carlo approach to estimate intraday VaR (IVaR) using tick-by-tick data. Employing a log ACD-ARMA-EGARCH model[3], they find that the approach produces reliable estimates of intra-day risk. The model benefits from its greater informational content than IVaR estimates based on regularly spaced data, and a greater flexibility with regard to the estimation time horizon. Recently, Huang (2010) uses an iterative Monte Carlo Simulation approach to produce a reliable VaR model, which more adequately models shocks to financial markets than a simple Monte Carlo Simulation. Further, Hoogerheide and van Dijk (2010) propose a technique for forecasting multiple step ahead VaR and ES using a Bayesian approach. Using data from the S&P 500 index, they find that the 10-day ahead forecasting for a single asset has similarities with the 1-day ahead forecasting (for a portfolio of 10 assets).

The present paper presents an empirical application of forecasting 1-step, 10-step and 20-step ahead 95% and 99% VaR and ES for 10 major worldwide stock indices[4]. Accurate VaR and ES forecasts are calculated by considering long memory within the conditional variance process and skewed Student-t conditionally distributed innovations. The Student-t distribution is commonly used in financial risk management (VaR models) with various

---

[3] In full, this is a log autoregressive conditional duration –autoregressive moving average – exponential GARCH model.

[4] At each point in time t, the risk forecasts for the τth day ahead is conducted; we do not sum up the forecasts made at time t for the next τ-day ahead daily variances.

methodologies being proposed (e.g. the two piece method by Hansen, 1994, Fernandez and Steel, 1998, Bauwens and Laurent, 2005, Azzalini and Capitanio, 2003, Zhu and Galbraith, 2010). In this paper, to fully capture not only the leptokurtosis but also the asymmetry of the portfolio returns, we incorporate the skewed version of the Student-*t* distribution proposed by Fernandez and Steel (1998). Further, according to Hoogerheide and van Dijk (2010), the model selection has an important effect on the numerical accuracy of the VaR and ES estimates.

The key contribution of this paper is to propose a new adaptation of the Monte Carlo simulation technique of Christoffersen (2003) for forecasting multiple step ahead VaR and ES. The present paper enables i) the incorporation of long memory in the volatility of the returns as well as ii) the utilization of skewed Student-*t* conditionally distributed innovations in estimating multi-period VaR and ES forecasts. At present, there are, to the best of the authors' knowledge, no studies within the literature, which estimate multi-period VaR or ES using either a fractionally integrated volatility model or leptokurtotic and asymmetric conditional distribution of innovations. Moreover, the proposed simulation-based algorithm differs from existing methods to produce long horizon VaR, in that it estimates the time path of volatility and density function for the returns and not just scaling the tail risk (see for example Wang et al. 2011).

The results show that the FIGARCH-skT model has a superior multi-period VaR and ES forecasting performance to the GARCH-skT model, for the 10-step and 20-step ahead time horizons. The result that accounting for fractional integration and asymmetric and leptokurtic conditional distribution improve the multi-period VaR and ES forecasting performance should prove to be valuable information for risk analysts and managers.

The remainder of the paper is organised as follows: Section 2 presents the framework of the GARCH-skT and FIGARCH-skT models. Section 3 shows the methods for modelling 1-step ahead and multiple step ahead VaR and ES, while Section 4 describes our data. Section 5 presents the empirical analysis of this paper and Section 6 concludes the paper and summarises the main findings.

2. Modelling GARCH-skT and FIGARCH-skT

To successfully capture the characteristics of financial returns data, many papers use GARCH family models under different distributional assumptions. In this paper, we assume

that the data generating process for the log-returns series, $\{y_t\}_{t=1}^T = \left\{\log\dfrac{p_t}{p_{t-1}}\right\}_{t=1}^T$, where p_t is

the closing price on trading day t, follows an ARCH process (Engle, 1982):

$$y_t = \mu_t + \varepsilon_t \,,$$
$$\varepsilon_t = \sigma_t z_t \,, \ \text{ where } \ z_t \sim f(0,1,\theta), \tag{1}$$
$$\sigma_t^2 = g(I_{t-1}; w).$$

The standardized error term z_t has a density function $f(.)$, where $E(z_t)=0$, $Var(z_t)=1$, and

θ represents the vector of parameters of f to be estimated. The conditional variance of the

error term, σ_t^2, is a time-varying, positive and measureable function $g(.)$ of the information

set (I_{t-1}) at time $t-1$, with w the vector of parameters to be estimated in the conditional

variance equation.

For the $\text{GARCH}(p,q)$ specification, $g(.)$ takes the functional form:

$$\sigma_t^2 = a_0 + \sum_{i=1}^q a_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \,. \tag{2}$$

Equation (2) can be rewritten with lag operators as follows:

$$\sigma_t^2 = a_0 + a(L)\varepsilon_t^2 + \beta(L)\sigma_t^2, \tag{3}$$

where $a(L)$ and $\beta(L)$ are lag operator polynomials of order q and p respectively.

Turning to the rate of decay of shocks to the conditional volatility process, Baillie et

al. (1996) noted that the distinction between integrated specifications, where shocks affect the

optimal volatility forecast indefinitely, as for example in the $\text{IGARCH}(p,q)$ specification

given by:

$$\Theta(L)(1-L)\varepsilon_t^2 = a_0 + (1 - \beta(L))(\varepsilon_t^2 - \sigma_t^2), \tag{4}$$

and covariance stationary models, where shocks to the volatility process decay exponentially,

such as the $\text{GARCH}(p,q)$ specification, was too sharp. To solve this, Baillie et al. (1996)

introduced the $\text{FIGARCH}(p,d,q)$ process, by replacing the first difference operator from

equation (4) with the fractional differencing operator $(1-L)^d$:

$$\Theta(L)(1-L)^d \varepsilon_t^2 = a_0 + (1 - \beta(L))(\varepsilon_t^2 - \sigma_t^2), \tag{5}$$

where $\Theta(L) = (1 - a(L) - \beta(L))(1 - L)^{-1}$. The roots of $\Theta(L)$ and $(1 - \beta(L))$ lie outside of the unit circle.

The fractional differencing operator $(1 - L)^d$ is defined as:

$$(1 - L)^d = \sum_{j=0}^{\infty} \pi_j L^j \quad \text{where} \quad \pi_j = \frac{\Gamma(j - d)}{\Gamma(j + 1)\Gamma(-d)}. \tag{6}$$

In the FIGARCH model, $0 < d < 1$ indicates that shocks to the conditional variance decay at a hyperbolic rate (Baillie et al., 1996). The FIGARCH model nests the $\text{IGARCH}(p, q)$ where $d = 1$, as well as the $\text{GARCH}(p, q)$, where $d = 0$. FIGARCH processes are strictly stationary and ergodic but are not weakly stationary since the second moment is infinite.

There is substantial evidence for the presence of long memory in volatility of daily and high frequency datasets; see Baillie et al. (1996) and Kilic (2011). Corsi (2009) shows that long memory specification improves the forecasting accuracy of realized volatility significantly. Moreover, recent evidence on volatility forecasting applied to high frequency datasets shows that "the forecasting accuracy is improved when the long memory property is taken into account"; see Chortareas et al. (2011)[5].

The conditional mean is modelled using an ARMA(1,0) whilst in the conditional variance it is assumed that $p = q = 1$. The fact that the values of time series are often taken to have been recorded at time intervals of one length when in fact they were recorded at time intervals of another, not necessarily regular, length is an effect known as the *non-synchronous trading effect*. Non-synchronous trading in the stocks making up an index induces autocorrelation in the return series. To control this Lo and MacKinlay (1988) suggested a first order autoregressive form for the returns' process. For more details see Campbell et al. 1997. Following Angelidis et al. (2004), we do not select the order of p and q according to a model selection criterion, such as the Akaike Information Criterion (AIC) or the Schwarz Bayesian Criterion (SBC).[6] They argue that in the majority of empirical studies the order of one lag has proven to work effectively in forecasting volatility for both GARCH and ARFIMA frameworks; hence, in this study we choose to set $p = q = 1$.

---

[5] Chortareas et al. (2011) argue that the FIGARCH model performs better than GARCH model when high frequency data on euro exchange rates are considered.
[6] According to Degiannakis and Xekalaki (2007, p.154) the commonly used in-sample methods of model selection such as AIC, SBC and Mean Squared Error (MSE), do not lead to the selection of a model that closely tracks future volatility. Moreover, Angelidis and Degiannakis (2007) provide evidence that an order of one lag has been shown to be sufficiently effective in modelling conditional volatility.

Further, since financial returns data is characterised by its skewness and its excess kurtosis, the standardised residuals are distributed on a skewed Student-*t* distribution, in preference to the normal distribution which, has been widely shown to underestimate risk, particularly in the tails of the returns distribution (see Giot and Laurent, 2003a, 2003b; Angelidis et al., 2004; Tang and Shieh, 2006; Kuester et al., 2006).

Therefore, the overall model is an AR(1)-FIGARCH$(1,d,1)$ with skewed Student-*t* distributed innovations, utilising the density function proposed by Fernandez and Steel (1998); see also Lambert and Laurent (2000, 2001)[7]:

$$y_t = c_0(1-c_1) + c_1 y_{t-1} + \varepsilon_t,$$

$$\varepsilon_t = \sigma_t z_t,$$

$$\sigma_t^2 = a_0 + (a_1 - b_1)\varepsilon_{t-1}^2 + \sum_{j=1}^{\infty}\left(\frac{d\Gamma(j-d)}{\Gamma(1-d)\Gamma(j+1)}L^j\left(\varepsilon_t^2 - a_1\varepsilon_{t-1}^2\right)\right) + b_1\sigma_{t-1}^2,$$

$$z_t \overset{i.i.d.}{\sim} skT(0,1;v,g),$$ (7)

$$f_{(skT)}(z_t;g,v) = \begin{cases} \dfrac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{\pi(v-2)}}\left(\dfrac{2s}{g+g^{-1}}\right)\left(1+\dfrac{sz_t+m}{v-2}g\right)^{-\frac{v+1}{2}} & if \quad z_t < -ms^{-1}, \\[4mm] \dfrac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{\pi(v-2)}}\left(\dfrac{2s}{g+g^{-1}}\right)\left(1+\dfrac{sz_t+m}{v-2}g^{-1}\right)^{-\frac{v+1}{2}} & if \quad z_t \geq -ms^{-1}, \end{cases}$$

where g and v are the asymmetry and tail parameters of the distribution, respectively, $m = \Gamma((v-1)/2)\sqrt{(v-2)}\left(\Gamma(v/2)\sqrt{\pi}\right)^{-1}\left(g - g^{-1}\right)$, and $s = \sqrt{g^2 + g^{-2} - m^2 - 1}$.

3. Modelling 1-step ahead and multiple step ahead VaR and ES

VaR at a given probability level $(1-\rho)$ is a single figure which represent a portfolio's worst possible outcome (either a significant loss when a long position is held, or an exceptionally high return if a short position is held), which is likely to occur under normal market conditions over a pre-determined period and for a given confidence level, i.e. $(1-\rho)$. However, the use of VaR has a number of limitations. There is no indication of the size of the loss when it exceeds the VaR figure. This problem can be overcome by calculating the ES of the portfolio which is a coherent risk measure[8]. In the event of a VaR violation, the ES is

---

[7] Note that AR(1) is presented as $y_t = c_0 + e_t, e_t = c_1 e_{t-1} + \varepsilon_t$, thus $(y_t - c_0) = c_1(y_{t-1} - c_0) + \varepsilon_t$.

[8] A risk measure ρ is coherent if it is in accordance with the properties of (i) sub-additivity, (ii) homogeneity, (iii) monotonicity and (iv) risk-free condition. These are described in the following equations: (i) $\rho(x) + \rho(y) \leq \rho(x+y)$; (ii) $\rho(tx) = t\rho(x)$; (iii) $\rho(x) \geq \rho(y)$ if $x \leq y$ and (iv) $\rho(x+n) = \rho(x) - n$. For further

defined as the conditional expected loss. Moreover, the majority of VaR models suffer from excessive VaR violations, implying an underestimation of market risk (Kuester et al., 2006)[9].

Having estimated the parameters of the model, the 1-step ahead VaR is calculated as:

$$VaR_{t+1|t}^{(1-\rho)} = \mu_{t+1|t} + F\left(\alpha;\theta^{(t)}\right)\sigma_{t+1|t}, \tag{8}$$

where $\mu_{t+1|t}$ and $\sigma_{t+1|t}$ are the conditional forecasts of the mean and of the standard deviation at time $t+1$, given the information available at time t, respectively. $F\left(\alpha;\theta^{(t)}\right)$ is the αth quantile of the assumed distribution, given the estimated parameters θ at time t.

The 1-step ahead ES forecast for long trading positions is the 1-day ahead expected value of the loss, given that the return at time $t+1$ falls below the corresponding value of the VaR forecast, and is defined by:

$$ES_{t+1|t}^{(1-\rho)} = E\left(y_{t+1} \mid \left(y_{t+1} \leq VaR_{t+1|t}^{(1-\rho)}\right)\right). \tag{9}$$

The proposed algorithm has been constructed in order to provide a methodological contribution to the multi-step VaR and ES forecasting under a fractionally integrated volatility framework for leptokurtic and asymmetrically distributed portfolio returns. The key innovation of this paper is the estimation of multiple step ahead VaR and ES for the FIGARCH-skT specification. The new methodology is based on the numerical technique presented in Xekalaki and Degiannakis (2010) and has been adapted from Christoffersen (2003).

Consequently, we suggest a new adaptation of Christoffersen (2003) method, for calculating multiple VaR and ES for FIGARCH-skT, using a number of steps arising from a new algorithm.

To generate the τ-step ahead VaR and ES forecasts for the AR(1)-FIGARCH$(1,d,1)$-skT model, set out in framework (7), we employ a Monte Carlo simulation technique. Steps .1 in the algorithm are required to produce leptokurtic and asymmetrically conditionally distributed log-returns. Since analytical expressions for the multi-period density are not available, Steps .2 and .3 are used for obtaining estimates for multiperiod VaR and ES based on the fractional differencing operator. The out-of-sample observations for each index, $\widetilde{T}$,

---

details, see Artzner et al. (1999). I.e, VaR is not sub-additive, which means that the VaR of an overall portfolio may be greater than the sum of the VaRs of its component parts.

[9] Our empirical results suggest that the method presented in this study tends to overestimate market risk for the 1-step ahead time horizon, but that this tendency diminishes for longer forecasting horizons, and particularly for the 10-step ahead forecasting period.

are divided into $\tilde{T}/\tau$ non-overlapping intervals of observations, with τ observations in each interval.[10] For each non-overlapping interval, we proceed as follows[11]:

Step 1.1: Generate random numbers $\{\breve{z}_{i,1}\}_{i=1}^{MC}$ from the skewed Student-t distribution, where $MC = 5{,}000$ denotes the number of draws (see Note 1 in the Appendix). The pseudo-random numbers are used to compute the innovations for period $t+1$ onwards.

Step 1.2: Create the hypothetical returns of time $t+1$, as (see Note 2 in the Appendix):

$$\breve{y}_{i,t+1} = \sigma_{t+1|t}\breve{z}_{i,1} + c_0^{(t)}\left(1 - c_1^{(t)}\right) + c_1^{(t)}y_t, \text{ for } i = 1,\ldots,MC. \tag{10}$$

The return at time $t+1$ is generated according to the AR(1) process. The value of the error term at time $t+1$, is simulated using the relation $\breve{\varepsilon}_{t+1} = \sigma_{t+1|t}\breve{z}_{i,1}$.

Step 1.3: Create the forecast variance for time $t+2$ as (see Note 3 in the Appendix):

$$\breve{\sigma}_{i,t+2}^2 = a_0^{(t)} + \left(a_1^{(t)} - b_1^{(t)}\right)\left(\sigma_{t+1|t}\breve{z}_{i,1}\right)^2 + d\left(\left(\sigma_{t+1|t}\breve{z}_{i,1}\right)^2 - a_1\varepsilon_{t/t}^2\right) +$$
$$\sum_{j=2}^{\infty}\left(\frac{d^{(t)}\Gamma\left(j - d^{(t)}\right)}{\Gamma\left(1 - d^{(t)}\right)\Gamma(j+1)}L^j\left(\varepsilon_{t+2|t+j}^2 - a_1^{(t)}\varepsilon_{t+1|t+j}^2\right)\right) + b_1^{(t)}\sigma_{t+1|t}^2. \tag{11}$$

The values of the innovations up to time t are extracted from the model's estimation, whilst the value of the innovation at time $t+1$ is estimated, as detailed above, and is treated separately.

Step 2.1: Generate further random numbers, $\{\breve{z}_{i,2}\}_{i=1}^{MC}$ from the skewed Student-t distribution, to be used to simulate the innovations for period $t+2$ onwards.

Step 2.2: Calculate the hypothetical returns of time $t+2$, using the AR(1) process,

$$\breve{y}_{i,t+2} = \breve{\sigma}_{i,t+2}\breve{z}_{i,2} + c_0^{(t)}\left(1 - c_1^{(t)}\right) + c_1^{(t)}\breve{y}_{i,t+1}, \text{ for } i = 1,\ldots,MC. \tag{12}$$

Step 2.3: Create the forecast variance for time $t+3$ as (see Note 4 in the Appendix):

$$\breve{\sigma}_{i,t+3}^2 = a_0^{(t)} + \left(a_1^{(t)} - b_1^{(t)}\right)\left(\breve{\sigma}_{i,t+2}\breve{z}_{i,2}\right)^2 + d\left(\left(\breve{\sigma}_{i,t+2}\breve{z}_{i,2}\right)^2 - a_1\left(\breve{\sigma}_{i,t+1}\breve{z}_{i,1}\right)^2\right) +$$
$$\frac{1}{2!}d(1-d)\left(\left(\breve{\sigma}_{i,t+1}\breve{z}_{i,1}\right)^2 - a_1\varepsilon_{t/t}^2\right) + \sum_{j=3}^{\infty}\left(\frac{d^{(t)}\Gamma\left(j - d^{(t)}\right)}{\Gamma\left(1 - d^{(t)}\right)\Gamma(j+1)}L^j\left(\varepsilon_{t+3|t+j}^2 - a_1^{(t)}\varepsilon_{t+2|t+j}^2\right)\right) + b_1^{(t)}\breve{\sigma}_{i,t+2}^2. \tag{13}$$

Innovation terms relating to periods $t+1$ onwards are simulated using the relation $\breve{\varepsilon}_{t+j} = \sigma_{t+j|t}\breve{z}_{i,j}$.

…

Repeat the process for **Step 3** through to **Step τ-1**.

…

---

[10] The use of non-overlapping intervals is necessary to avoid autocorrelation in the forecast errors.
[11] The program code for this simulation is available from the authors on request.

Step τ.1: Generate further random numbers, $\left\{\breve{z}_{i,\tau}\right\}_{i=1}^{MC}$, from the skewed Student-*t* distribution, to be used to simulate the innovations for period $t+\tau$ onwards.

Step τ.2: Calculate the hypothetical returns of time $t+\tau$,

$$\breve{y}_{i,t+\tau} = \breve{\sigma}_{i,t+\tau}\breve{z}_{i,\tau} + c_0^{(t)}\left(1-c_1^{(t)}\right) + c_1^{(t)}\breve{y}_{i,t+\tau-1}.$$ (14)

Step τ.3: Construct a density function for the returns at time $t+\tau$ using the 5,000 simulated returns. Calculate the τ-day $(1-\rho)$VaR figure for the left-hand tail of this distribution, i.e.

$$VaR_{t+\tau|t}^{(1-\rho)} = f_\alpha\left(\left\{\breve{y}_{i,t+\tau}\right\}_{i=1}^{MC}\right).$$

Repeat **Steps 1.1** through to τ.3 for each of the non-overlapping intervals of τ observations, such that a total of $\tilde{T}/\tau$ VaR forecasts will be produced.

The τ-day ahead ES forecast for long trading positions is the τ-day ahead expected value of the loss, given that the return at time $t+\tau$ falls below the corresponding value of the VaR forecast, and is defined by:

$$ES_{t+\tau|t}^{(1-\rho)} = E\left(y_{t+\tau} \mid \left(y_{t+\tau} \leq VaR_{t+\tau|t}^{(1-\rho)}\right)\right),$$ (15)

The value of the τ-day ahead ES measure is given by:

$$ES_{t+\tau|t}^{(1-\rho)} = E\left(VaR_{t+\tau|t}^{(1-\tilde{\rho})}\right), \quad \forall 0 < \tilde{\rho} < \rho.$$ (16)

Following Dowd (2002), to calculate the ES we divide the tail of the probability distribution of returns into a large number $\tilde{k}$ of slices each with identical probability mass, estimate the τ-day ahead VaR attached to each slice and find the mean of these VaRs to estimate the τ-day ahead ES[12].

$$ES_{t+\tau|t}^{(1-\rho)} = \tilde{k}^{-1}\sum_{i=1}^{\tilde{k}}\left(VaR_{t+\tau|t}^{\left(1-\rho+i\rho(\tilde{k}+1)^{-1}\right)}\right).$$ (17)

4. Data Description

In order to examine the robustness of the VaR and ES forecasts produced by the proposed forecasting mode, VaR and ES forecasts were generated using daily returns data from 10 developed market stock indices. The indices are: Austrian Traded Index (ATXINDX), French Cotation Assistée en Continu - Continuous Assisted Quotation 40 (FRCAC40), Deutscher Aktien IndeX - Dax 30 Performance (DAXINDX), UK Financial

---

[12] In this study we take $\tilde{k}$ to be 5,000.

Times Stock Exchange 100 (FTSE100), Hang Seng (HNGKNGI), Nikkei 225 Stock Average (JAPDOWA), OMX Stockholm (SWSEALI), National Association of Securities Dealers Automated Quotations - NASDAQ 100 (NASA100), NYSE Composite (NYSEALL) and Standard and Poor's 500 Composite (S&PCOMP). Our sample considers data from major world stock market indices with the longest continuous history[13]. The selection is based on the following criteria: i) the indices' market capitalization, and ii) the fact that they are the most publicly quoted stock market indices. In addition, most indices are considered as benchmark indices for large stocks (e.g. NASDAQ, NYSE, S&P500) traded internationally as they contain about 70%-80% of the value of their individual stocks (the selected indices track the performance of large companies based in the specific country). Moreover, the list includes world's top stock exchanges by value shares traded as reported by World Federation of Exchanges Industry Association (WFE); see www.world-exchanges.org. The data, which was obtained from *Datastream*® for the period from 12th January, 1989 until 12th February, 2009, was conditioned to remove any non-trading days. Thus, the total number of log returns for a given index, T, ranged from 4,924 for the Nikkei 225 Stock Average, to 5,051 for the FTSE100 Index. Based on a rolling sample of $\hat{T} = 2,000$ observations, a total of $\tilde{T} = T - \hat{T}$ out-of-sample forecasts were produced for each model (with the parameters of the conditional mean, conditional variance and density function re-estimated each trading day) [14].

Descriptive statistics for the daily log-returns for the selected indices are given in Table 1. All of the returns distributions are leptokurtic and the majority are negatively skewed. The Jarque-Bera test results indicate that none of the log-returns series follow a Gaussian distribution. The absolute value of the log returns are significantly positively autocorrelated for a high number of lags. Examining the correlograms for the various indices, the decay in the value of the autocorrelation coefficients is initially rapid, before slowing and is suggestive of the hyperbolic decay which is typical for a long memory volatility process[15].

<Insert Table 1 about here>

Table 2 reports the full sample parameter estimates of the FIGARCH-skT model in order to provide a fair amount of evidence that i) the long-memory of conditional volatility as well as ii) a leptokurtic and asymmetric conditional distribution of innovations are present. The long memory parameter is statistically significant for all the indices supporting the

---

[13] For example, the Austrian ATX as well as the Swedish OMX indices are the leading indices of the Wiener Borse and the Stockholm Stock Exchange, respectively; both are two of the world's oldest exchanges (Wiener Borse was founded in 1771, while Stockholm Exchange was founded in 1863).

[14] The estimations were carried out using the G@RCH 6.0 (Laurent, 2009) package of Ox (Doornik, 2009).

[15] Correlograms for the absolute log returns of the 10 indices are available from the authors on request.

presence of long-memory. Moreover, the parameters of the skewed Student-t distribution are statistically significant strongly supporting the use of a skewed and leptokurtic distribution.

<Insert Table 2 about here>

5. Empirical Analysis

5.1. Evaluation Framework

A model is considered to accurately forecast the τ-step ahead VaR if it cannot be rejected by both the independence and conditional coverage hypotheses. Potential clustering of the VaR violations is an important consideration, and is tested for using the independence hypothesis. Christoffersen (1998) examines whether the instances of VaR failure are independent, based on the likelihood ratio statistic given below:

$$LR_{IN} = 2\log\left((1-\pi_{01})^{n_{00}} \pi_{01}{}^{n_{01}} (1-\pi_{11})^{n_{10}} \pi_{11}{}^{n_{11}}\right) - 2\log\left(\left(1-\frac{N}{\tilde{T}}\right)^{n_{00}+n_{10}} \left(\frac{N}{\tilde{T}}\right)^{n_{01}+n_{11}}\right) \sim \chi_1^2, \quad (18)$$

where, N is the number of days on which a violation occurred across the total VaR estimation period $\tilde{T}$, n_{ij} is the number of observations with value i followed by j, for $i,j = 0,1$, and $\pi_{ij} = \dfrac{n_{ij}}{\sum_j n_{ij}}$ are the corresponding probabilities[16]. The purpose of the test is to examine the null hypothesis that the VaR failures are independent and are spread over the whole estimation period, against the alternative hypothesis that the failures tend to be clustered. The main advantage of the test is that it can reject a model that generates either too many or too few clustered exemptions (see Cheng and Hung, 2010). As Angelidis et al. (2004) argue, the Christoffersen (2003) procedure can be used to separate clustering effects from distributional assumption effects.

The conditional coverage hypothesis (Christoffersen, 1998, 2003) combines Kupiec's (1995) test with independence hypothesis, and examines the null hypothesis that the observed violation rate, $N/\tilde{T}$, is statistically equal to the expected violation rate, ρ as well as that the VaR failures are independently distributed over time. In order to test this null hypothesis, the likelihood ratio statistic is:

$$LR_{cc} = -2\log\left((1-\rho)^{\tilde{T}-N} \rho^N\right) + 2\log\left((1-\pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1-\pi_{11})^{n_{10}} \pi_{11}^{n_{11}}\right) \sim \chi_2^2. \quad (19)$$

---

[16] $i,j = 1$ indicates that a violation has occurred, whereas $i,j = 0$ indicates the converse. π_{ij} indicates the probability that $j(=0,1)$ occurs at time t, given that $i(=0,1)$ occurred at time $t-1$. The null hypothesis is $H_0 : \pi_{01} = \pi_{11}$, which is tested against the alternative $H_1 : \pi_{01} \neq \pi_{11}$.

The most widely applied tests are those of Kupiec (1995) and Christoffersen (2003). However, testing for the validity of the VaR forecasts other tests have also been considered, i.e. density forecasts test of Berkowitz (2001), CAViaR test of Engle and Manganelli (2004), and they provide qualitatively similar results.

If the null hypothesis of both the independence and conditional hypotheses is not rejected for a particular model, then we conclude that the model produces the expected proportion of VaR violations, and that these violations are not clustered together.

However, it does not provide a method for distinguishing between the performances of the various models for which this is the case. Sarma et al. (2003) and Angelidis and Degiannakis (2007) suggest a two-stage backtesting procedure. In the first stage, the VaR forecasting ability of the candidate models is investigated, and in the second stage the forecasting accuracy of the models, which are judged to forecast the VaR adequately in the first stage, is compared. For the present study, in the first stage, the VaR forecasting ability of the candidate models is investigated according to the likelihood ration statistics in (18) and (19). For the second stage, the mean squared error, or $MSE = \tau \widetilde{T}^{-1} \sum \psi_{t+\tau}$, is calculated for the loss function:

$$\psi_{t+\tau} = \begin{cases} \left(y_{t+\tau} - ES_{t+\tau|t}^{(1-\rho)}\right)^2 & if \ y_{t+\tau} < VaR_{t+\tau|t}^{(1-\rho)}, \\ 0 & otherwise. \end{cases} \tag{20}$$

Therefore, we pick risk models that calculate the VaR accurately, as the prerequisite of independence and correct conditional coverage is satisfied, and provide more precise ES forecasts, as they minimize the *MSE*. The *MSE* evaluates the τ-day ahead ES forecasts. Thus for each VaR failure we compare the actual return to the forecasted return, given that the VaR is violated. Hence, the model will be deemed to perform well if:

i)   the VaR failures occur independently of each other (Christoffersen, 1998 test);
ii)   the observed failure rate equals the expected failure rate (Christoffersen, 1998, 2003 test);
iii)   the *MSE* based on the quadratic loss between the actual and expected returns in the event of a VaR violation is minimised (Hansen, 2005 test).

Finally, we should check whether the *MSE* values of the models differ statistically significant. Hansen (2005) proposed the Superior Predictive Ability, or SPA, test for comparing the performances of two or more forecasting models in terms of a predefined loss function, $\psi_{t+\tau}$. Let us denote as $\psi_{t+\tau}^{(A)}$ and $\psi_{t+\tau}^{(B)}$ the ES loss function for the FIGARCH-skT

and GARCH-skT models, respectively. The null hypothesis that the FIGARCH-skT model is not outperformed by the GARCH-skT model, or $H_0 : E\left(\psi_{t+\tau}^{(A)} - \psi_{t+\tau}^{(B)}\right) \leq 0$, is tested against the alternative, $H_1 : E\left(\psi_{t+\tau}^{(A)} - \psi_{t+\tau}^{(B)}\right) > 0$.[17] Obviously, the $\psi_{t+\tau}^{(A)}$ and $\psi_{t+\tau}^{(B)}$ may be considered for various modifications of the loss function, i.e. the absolute distance, $\left|y_{t+\tau} - ES_{t+\tau|t}^{(1-p)}\right|$, as well as the absolute percentage distance, $\left|\left(y_{t+\tau} - ES_{t+\tau|t}^{(1-p)}\right)\middle/ y_{t+\tau}\right|$.

5.2. Empirical Results

The empirical results for forecasting 1-step, 10-step and 20-step ahead 95% VaR using the FIGARCH and GARCH models based on skewed Student-t distribution are shown in Tables 3 and 4, respectively. According to the conditional coverage test, the FIGARCH-skT model produces an adequate forecasting performance for 5, 10 and 10 indices (out of the 10 indices tested) at the 1-step ahead, 10-step ahead and 20-step ahead forecasting horizons, respectively. This compares favourably to the GARCH-skT model which produces an adequate forecasting performance for 4, 8 and 10 indices (out of the 10 indices tested) at the 1-step ahead, 10-step ahead and 20-step ahead forecasting horizons, respectively. Accounting for fractional integration in the volatility model improves the adequacy of the 95% VaR forecasts at the 10-step horizon, for which the FIGARCH-skT model adequately predicts losses for all the 10 indices.

For the FIGARCH-skT model, the results of the independence test indicate that the null hypothesis that the VaR violations occur independently cannot be rejected for any of the indices for any time horizon. Similarly, for the GARCH model, the null hypothesis of independence between the VaR violations cannot be rejected for any of the indices at any time horizon except for the ATXINDX at the 10-step ahead time period. Thus, the FIGARCH-skT and GARCH-skT specifications demonstrate a comparable performance in terms of the independence of the VaR violations, in the multi-period VaR forecasts.

<Insert Table 3 about here>

<Insert Table 4 about here>

Furthermore, the average 95% ES figure tends to decrease as the forecasting horizon increases for the FIGARCH-skT model across all of the indices. By contrast, the average 95% ES figure for the GARCH-skT model decreases between the 1-step and 10-step ahead

---

[17] The SPA statistic equals to $SPA = \tilde{T}^{-1} \sum_{t=1}^{\tilde{T}/\tau} \left(\psi_{t-\tau}^{(A)} - \psi_{t-\tau}^{(B)}\right) \left(V\left(\tau \tilde{T}^{-1} \sum_{t=1}^{\tilde{T}/\tau} \left(\psi_{t-\tau}^{(A)} - \psi_{t-\tau}^{(B)}\right)\right)\right)^{-1/2}$.

forecasting horizons for all of the indices, but, subsequently, increases between the 10-step and 20-step ahead forecasting horizons. The *MSE* of the ES figures for the FIGARCH-skT model are generally lower than those for the GARCH-skT model, especially as the forecasting horizon increases. For example, the *MSE* of the 1-step ahead 95% ES figures are greater for the FIGARCH-skT model than the GARCH-skT model for 5 of the 10 indices. However, the *MSE* of the 10-step ahead ES figures for the FIGARCH-skT model are lower than, or equal to, those for the GARCH-skT model for all 10 indices. Furthermore, the *MSE* of the 20-step ahead ES figures are lower for the FIGARCH-skT model than the GARCH-skT model for 8 of the 10 indices.

In order to evaluate the performance of the models, we proceed to the statistical comparison of the mean squared error loss function in (20). Table 5 presents the *p*-values of the SPA test for the null hypothesis that the FIGARCH-skT model outperforms the GARCH-skT model. A high *p*-value indicates evidence in support of the hypothesis that the FIGARCH-skT model is superior to the GARCH-skT. As all the *p*-values of the SPA test are greater than 0.05, there is evidence suggesting that the FIGARCH-skT model does not demonstrate an inferior performance to the GARCH-skT model in forecasting losses when the VaR figure is breached. In the second part of Table 5, the mean absolute error loss function is applied, which provides qualitatively similar results[18]. We, therefore, conclude that the FIGARCH-skT model does demonstrate a superior performance to the GARCH-skT model in forecasting multiple step ahead losses.

<center><Insert Table 5 about here></center>

The empirical results for forecasting 1-step, 10-step and 20-step ahead 99% VaR using the FIGARCH and GARCH models based on skewed Student-*t* distribution are shown in Tables 6 and 7, respectively. The results are qualitatively similar to the 95% case; the fractional integration in the volatility model improves the adequacy of the 99% VaR forecasts at the 10-step and 20-step horizons. The conditional coverage test informs us that the FIGARCH-skT model produces adequate forecasting performance for all the 10 indices at the 10-step ahead and 20-step ahead forecasting horizons. This compares favourably to the GARCH-skT model which produces an adequate forecasting performance for 8 and 9 indices (out of the 10 indices tested) at the 10-step ahead and 20-step ahead forecasting horizons, respectively. On the contrary, both models fail to fulfil the conditional coverage for the next

---

[18] The $\psi_{t+\tau}^{(A)}$ and $\psi_{t+\tau}^{(B)}$ have been computed for $\psi_{t+\tau} = \left| y_{t+\tau} - ES_{t+\tau|t}^{(1-\rho)} \right|$ if $y_{t+\tau} < \text{VaR}_{t+\tau|t}^{(1-\rho)}$ and $\psi_{t+\tau} = 0$ otherwise. The mean absolute percentage error loss function was also applied and provided similar results.

trading day; in 8(5) out of the 10 indices the FIGARCH(GARCH) model does not produce an adequate forecasting performance at the 1-step ahead time horizon.

The *MSE* of the 10-step ahead 99% ES figures for the FIGARCH-skT model are lower than those for the GARCH-skT model for 9 out of the 10 indices. As in the case of 95% confidence interval, the evaluation of the performance of the models with the SPA test provide evidence in support of a superior performance to the GARCH-skT model in forecasting multiple step ahead losses[19].

<center><Insert Table 6 about here></center>
<center><Insert Table 7 about here></center>

The magnitude of the observed failure rate for each forecasting horizon suggests that both models tend to over-forecast the VaR figure at the 1-step ahead time horizon, but that this tendency diminishes for longer forecasting horizons, and in particular for the 10-step ahead forecasting horizon. This contrasts with the findings of Kuester et al. (2006) who found that VaR models tend to underestimate the true VaR figure for the 1-step ahead time horizon.

It is interesting to compare the volatility of the returns. Figures 1 and 2 show plots for observed returns against the 95% VaR forecasts resulting from the FIGARCH-skT and GARCH-skT models for the SWSEALI and HNGKNGI indices for the 1-step, 10-step and 20-step ahead forecasting horizons[20]. Looking at the plots for the SWSEALI index (Figure 1a), it appears that although the most recent returns (representing the start of the global financial crisis) are quite volatile in the 1-step 95% VaR plot, the graph (Figure 1b) showing every 10th return (for the 10-step VaR forecasts) displays less volatility. Comparing this to the plots for the HNGKNGI index (Figure 2), which is slightly more volatile, when we consider all the returns, than the SWSEALI index, but is much more volatile when we consider every 10th return figure (for the 10-step VaR forecasts). The proposed method performs particularly well for the SWSEALI index.

<center><Insert Figure 1 about here></center>
<center><Insert Figure 2 about here></center>

5.3. Square-root Rule

As noted by Danielsson and Zigrand (2006) and Engle (2004), the square root of time scaling rule appears to lead to inadequate VaR forecasts and, hence, the objective of the Basle Committee is not addressed satisfactorily. Table 8 reports the forecasting 10-day-ahead 95%

---

[19] The p-values of the SPA test are available to the readers upon request.
[20] Corresponding figures for the remaining indices are available from the authors on request.

VaR and 95% ES using the FIGARCH-skT model based on the square root rule[21]. According to the independence and conditional coverage tests, at the 10-step ahead forecast, the square of root rule does not produce adequate risk forecasts for just 1 out of the 10 indices. Even the square-root rule under an appropriate modelling framework is able to provide proper risk forecasts. However, the comparison of the square-root rule to the FIGARCH-skT model under the Monte Carlo simulation technique is in favour of the proposed method. There is significant difference of the percentage of observed violations between the square-root rule (in Table 8) and the proposed simulation technique. For 9 out of the 10 indices, the observed exception rate is closer to 5% for the proposed simulation technique compared to the square-root rule (for the DAXINDX the absolute difference of observed exception rate from 5% is 0.26% under the square-root rule and 0.92% under the proposed simulation technique).

Therefore, we provide a further support that the proposed multi-period forecasting risk method is not just a byproduct of a shift in the unconditional variance but instead it does capture the long-memory of volatility. Hence, our findings are in accordance with Danielsson and Zigrand (2006); "even if the square root of time rule has widespread applications in the Basel Accords, it fails to address the objective of the Accords".

<center><Insert Table 8 about here></center>

6. Conclusion and Suggestions for Further Research

This paper has introduced a new adaptation to the FIGARCH-skT model of the Monte Carlo simulation technique of Christoffersen (2003) for forecasting multiple step ahead 95% and 99% VaR and ES. Much of the existing literature on VaR forecasting is limited to the 1-step ahead horizon, or employs unjustifiable assumptions to produce multiple step ahead forecasts using scaling rules such as the square root of time rule.

The VaR forecasting accuracy of the simulation technique was tested on 10 worldwide stock indices. Based on a two-stage backtesting procedure, the VaR forecasting ability of the candidate models is investigated, and the forecasting accuracy of the models, which are judged to forecast the VaR adequately, is compared. The Superior Predictive Ability test compares the forecasting performance of the competing models in terms of a predefined loss function. The results show that the FIGARCH-skT model has a superior 95% and 99% VaR and ES forecasting performance to the GARCH-skT model, for the 10-step and

---

[21] The square-root rule assumes that τ-day ahead variance forecasts are equal for each day τ. Therefore, under the square-root rule, the 10-day ahead VaR is constructed as $VaR_{t+10|t}^{(95\%)} = \mu_{t+1|t} + F(\alpha; \theta^{(t)}) \sigma_{t+1|t}$.

20-step ahead time horizon. Furthermore, the tendency for the models to over predict the VaR figure for the 1-step ahead horizon, appeared to diminish for longer forecasting intervals.

The Basel regulations require a 10-day VaR. The FIGARCH-skT model performs accurately at the 10-day and 20-day ahead 95% and 99% VaR. Is there anything in the structure of the model or the nature of the markets that may cause this to happen? For the 1-day horizon, the long memory structured model does not perform better that the short memory. For the 10-day horizon, the FIGARCH-skT model appeared to produce its best forecasts, providing evidence that the superiority of the long memory volatility modelling is detected in two-weeks (in calendar time) forecasts, as the Basel regulations require. Although, the 20-trading-day (or the one-month in calendar time) horizon is considered a faraway point in time to be predicted, the FIGARCH-skT model provides accurate 95% and 99% VaR forecasts. However, for the 20-day time horizon, the results of the conditional coverage tests were highly sensitive to the number of VaR violations such that a very small number of additional (or fewer) violations can be pivotal in determining whether or not the forecasting performance of the model is deemed to be adequate.

Considering the case of ES forecasting, the FIGARCH-skT model does demonstrate a superior performance to the GARCH-skT model in forecasting losses given that a VaR violation has occurred. The 10-step ahead quadratic loss between the actual and expected returns in the event of a 95% VaR violation is lower for the long memory volatility model for all 10 indices. The 10-step ahead *MSE* in the event of a 99% VaR violation is lower for the long memory volatility model for 9 indices. Furthermore, the 20-step ahead loss function for the FIGARCH-skT model is lower for 9 of the 10 indices. Since the findings suggest that FIGARCH-skT models have a superior 95% and 99% VaR and ES forecasting performance, for the 10-step and 20-step ahead time horizons, risk managers and analysts should apply our technique to obtain accurate 95% and 99% risk forecasts.

Future research would require the use of other type of financial time series (i.e. exchange rates, stocks, futures, portfolio of assets etc.), as well as of different frequency datasets (e.g. intra-day data), for testing the performance of the proposed method. A further investigation of the trade-off between complexity and accuracy of VaR and ES forecasting would be the expansion of the FIGARCH-skT model to an asymmetric conditional volatility framework such as the Tse's (1998) FIAPARCH, Davidson's (2004) HYGARCH model, or indeed both. The multi-period forecasting performance of regime-depended GARCH models, which are well established to show that financial volatility behaves with asymmetrical,

nonlinear, and regime-switching dynamics (see Haas et al., 2004 and Huang, 2011), may also be investigated in future research.

Acknowledgement

References

Alexander, C. 2008. *Value-at-Risk Models*. New York: John Wiley & Sons, Inc.

Angelidis, T. and S. Degiannakis. 2007. Backtesting VaR Models: A Two-Stage Procedure. *Journal of Risk Model Validation* 1: 2: 1-22.

Angelidis, T., A. Benos and S. Degiannakis. 2004. The use of GARCH models in VaR estimation. *Statistical Methodology* 1: 105-128.

Artzner, P., F. Delbaen, J. Eber and D. Heath. 1999. Coherent Measures of Risk. *Mathematical Finance,* 9: 3: 203-228.

Azzalini, A. and A. Capitanio. 2003. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew *t* distribution, *Journal of the Royal Statistical Society B* 65**:** 2: 367–389.

Baillie, R., T. Bollerslev and H. Mikkelsen. 1996. Fractionally integrated generalised autoregressive conditional heteroscedasticity. *Journal of Econometrics* 74: 1: 3-30.

Basel Committee on Banking Supervision. 1988. *International Convergence of Capital Measurement and Capital Standards*. Basel.

Basel Committee on Banking Supervision. 1996*. Amendment to the Capital Accord to incorporate Market Risks.* Basel.

Basel Committee on Banking Supervision. 2006. *International Convergence of Capital Measurement and Capital Standards – A Revised Framework*. Basel.

Basel Committee on Banking Supervision. 2009. *Revisions to the Basel II Market Risk Framework*. Basel.

Bauwens, L. and S. Laurent. 2005. A new class of multivariate skew densities, with application to GARCH models. *Journal of Business and Economic Statistics* 23: 346–354.

Beltratti, A. and C. Morana. 1999. Computing value at risk with high frequency data. *Journal of Empirical Finance* 6: 5*:* 431-444.

Berkowitz, J. 2001. Testing Density Forecasts, with Applications to Risk Management. *Journal of Business and Economic Statistics*, 19: 465-474.

Berkowitz, J. and J. O'Brien. 2002. How Accurate Are Value-at-Risk Models at commercial Banks? *The Journal of Finance*, 57: 1093–1111.

Bollerslev, T. 1986. Generalised autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31**:** 3: 307-327.

Brooks, C. and G. Persand. 2003. Volatility Forecasting for Risk Management. *Journal of Forecasting* 22**:** 1: 1-22.

Campbell, J., A. Lo and A.C. MacKinlay. 1997. *The Econometrics of Financial Markets*. New Jersey: Princeton University Press.

Caporin, M. 2008. Evaluating Value-at-Risk measures in presence of long memory conditional volatility, *Journal of Risk* 10: 3: 79-110.

Cheng, W.H. and J.C. Hung. 2010. Skewness and leptokurtosis in GARCH-typed VaR estimation of petroleum and metal asset returns. *Journal of Empirical Finance* 18: 1: 160-173.

Chortareas, G., Y. Jiang and J. Nankervis. 2011. Forecasting exchange rate volatility using high-frequency data: is the euro different? *International Journal of Forecasting* 27: 1089–1107.

Christoffersen, P. 1998. Evaluating interval forecasts. International Economic Review, 39: 841-862.

Christoffersen, P. 2003. *Elements of Financial Risk Management.* USA: Elsevier Science.

Corsi, F. 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7: 2: 174–196.

Crouhy, M., D. Galai and R. Mark. 2001. *Risk Management*. New York: McGraw-Hill.

Danielsson, J. 2002. The emperor has no clothes: Limits to risk modelling. *Journal of Banking and Finance* 26: 7: 1273-1296.

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: 10: 2701–2713.

Davidson, J. 2004. Moment and Memory Properties of Linear Conditional Heteroscedasticity Models, and a New Model. *Journal of Business and Economic Statistics*, 22: 1: 16-29.

Degiannakis, S. and E. Xekalaki. 2007. Assessing the Performance of a Prediction Error Criterion Model Selection Algorithm in the Context of ARCH Models. *Applied Financial Economics* 17: 149-171.

Dionne, G., P. Duchesne and M. Pacurar. 2009. Intraday Value at Risk (IVaR) using tick-by-tick data with application to the Toronto Stock Exchange. *Journal of Empirical Finance* 16: 777-792.

Doornik, J.A. 2009. Ox: Object Oriented Matrix Programming, 6.0. London: Timberlake Consultants Press.

Dowd, K. 2002. *Measuring Market Risk.* New York: John Wiley & Sons, Inc.

Drost, F.C. and T.E. Nijman. 1993. Temporal aggregation of GARCH processes. *Econometrica* 61: 909-927.

Engle, R.F. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica* 50: 987-1008.

Engle, R.F. 2004. Risk and Volatility: Econometric Models and Financial Practice. *The American Economic Review* 94: 3: 405-420.

Engle, R.F. and S. Manganelli. 2004. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business and Economic Statistics* 22: 4: 367-381.

Fernandez, C. and M.F.J. Steel. 1998. On Bayesian modelling of fat tails and skewness, *Journal of the American Statistical Association* 93: 359–371.

Giot, P. and S. Laurent. 2003a. Value-at-Risk for Long and Short Trading Positions. *Journal of Applied Econometrics* 18: 6: 641-663.

Giot, P. and S. Laurent. 2003b. Market risk in commodity markets: a VaR approach. *Energy Economics* 25**:** 5: 435-457.

Giot, P. and S. Laurent. 2004. Modelling daily Value-at-Risk using realised volatility and ARCH type models. *Journal of Empirical Finance* 11: 3: 379-398.

González-Rivera, G., T.H. Lee and S. Mishra. 2004. Forecasting volatility: a reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. *International Journal of Forecasting* 20: 629-645.

Grané, A. and H. Veiga. 2008. Accurate minimum capital risk requirements: a comparison of several approaches. *Journal of Banking and Finance* 32: 11: 2482–2492.

Haas, M., S. Mittnik, and M.S. Paolella. 2004. A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics* 2: 493–530.

Hansen, B.E. 1994. Autoregressive conditional density estimation. *International Economic Review* 35: 705–730.

Hansen, P.R. 2005. A Test for Superior Predictive Ability. *Journal of Business and Economic Statistics* 23: 365-380.

Hartz, C., S. Mittnik and M. Paolella. 2006. Accurate value-at-risk forecasting based on the normal-GARCH model. *Computational Statistics and Data Analysis* 51: 4: 2295-2312.

Hoogerheide, L. and H.K. van Dijk. 2010. Bayesian forecasting of Value at Risk and Expected Shortfall using adaptive importance sampling. *International Journal of Forecasting* 26: 231-247.

Huang, A. 2010. An optimization process in Value-at-Risk estimation. *Review of Financial Economics* 19: 3: 109-116.

Huang, A. 2011. Volatility modeling by asymmetrical quadratic effect with diminishing marginal impact. *Computational Economics* 37: 301-330.

Kilic, R. 2011. Long memory and nonlinearity in conditional variances: A smooth transition FIGARCH model. *Journal of Empirical Finance* 18: 2: 368-378.

Kinateder, H. and N. Wagner. 2010. Market Risk Prediction under Long Memory: When VaR is higher than expected. *Passau University,* Working Paper.

Kuester, K., S. Mittnik and M.S. Paolella. 2006. Value-at-Risk prediction: a comparison of alternative strategies. *Journal of Financial Econometrics* 4: 1: 53-89.

Kupiec, P. 1995. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives* 3: 2: 73-84.

Laurent, S. 2009. G@RCH 6, Estimating and Forecasting ARCH Models. London: Timberlake Consultants Press.

Lambert, P. and S. Laurent. 2000. Modeling Skewness Dynamics in Series of Financial Data. Discussion Paper, Institut de Statistique, Louvain-la-Neuve.

Lambert, P. and S. Laurent. 2001. Modeling Financial Time Series Using GARCH-Type Models and a Skewed Student Density.Universite de Liege, Mimeo.

Lo, A. and A.C. MacKinlay. 1988. Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test. *Review of Financial Studies* 1: 41-66.

McMillan, D. and D. Kambouroudis. 2009. Are Riskmetrics forecasts good enough? Evidence from 31 stock markets. *International Review of Financial Analysis* 18: 3: 117-124.

Sarma, M., S., Thomas and A., Shah. 2003. Selection of VaR models. *Journal of Forecasting* 22: 4: 337-358.

Semenov, A. 2009. Risk factor beta conditional Value-at-Risk. *Journal of Forecasting* 28: 6: 549-558.

Shao, X.D., Y.J. Lian and L.Q. Yin. 2009. Forecasting Value-at-Risk using high frequency data: The realized range model. *Global Finance Journal* 20: 2: 128-136.

Tang, T. and S. Shieh. 2006. Long memory in stock index futures markets: A value-at-risk approach. *Physica A* 366: 437-448.

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

Wang, J., J. Yeh and N. Cheng. 2011. How accurate is the square-root-of-time rule in scaling tail risk: A global study. *Journal of Banking and Finance* 35: 1158-1169.

Xekalaki, E. and S. Degiannakis. 2010. *ARCH Models for Financial Applications*. New York: John Wiley & Sons, Inc.

Zhu, D. and J.W. Galbraith. 2010. A generalized asymmetric Student-*t* distribution with application to financial econometrics. *Journal of Econometrics* 157: 2: 297-305.

Appendix

Note 1: We define the scheme follows in order to create random draws from the skewed Student-*t* distribution based on Lambert and Laurent (2001).

1. Generate random numbers $\{\rho_i\}_{i=1}^{MC}$ from the standard uniform distribution, where *MC* denotes the number of draws[22].

2. For each ρ_i compute the $\breve{z}_{i,1}$ random draw as[23]:

$$F_{(skT)}^{-1}(\rho_i; v, g) = \begin{cases} \dfrac{g^{-1}F_{(T)}^{-1}\left(\dfrac{\rho_i}{2}\left(1+g^2\right); v\right) - m}{s} & if \quad \rho_i < \dfrac{1}{1+g^2} \\[6mm] \dfrac{-gF_{(T)}^{-1}\left(\dfrac{1-\rho_i}{2}\left(1+g^{-2}\right); v\right) - m}{s} & if \quad \rho_i \geq \dfrac{1}{1+g^2} \end{cases}, \qquad (A1)$$

where $F_{(T)}^{-1}(\rho_i; v)$ corresponds to the inverse CDF of the Student-*t* distribution with v degrees of freedom.

The $F_{(skT)}^{-1}(\rho_i; v, g)$ corresponds to the inverse CDF of the skewed Student-*t* distribution with g and v denoting the asymmetry and tail parameters of the distribution, respectively.

---

[22] We adopt Christoffersen's (2003) symbol (MC) for the number of draws.
[23] $\breve{z}_{i,1} \equiv F_{(skT)}^{-1}(\rho_i; v, g)$.

Equation (A1) relates the inverse CDF of the skewed Student-t with the inverse CDF of the symmetric Student-t.

Note 2: The 1-step ahead variance forecast is computed as:

$$\sigma_{t+1|t}^2 = a_0^{(t)} + \left(a_1^{(t)} - b_1^{(t)}\right)\varepsilon_{t|t}^2 + \sum_{j=1}^{\infty}\left(\frac{d^{(t)}\Gamma\left(j-d^{(t)}\right)}{\Gamma\left(1-d^{(t)}\right)\Gamma(j+1)}L^j\left(\varepsilon_{t+1|t+j}^2 - a_1^{(t)}\varepsilon_{t|t+j}^2\right)\right) + b_1^{(t)}\sigma_{t|t}^2 .$$

Note 3: This is based on the 2-step ahead volatility forecasting formula:

$$\breve{\sigma}_{i,t+2}^2 = a_0^{(t)} + \left(a_1^{(t)} - b_1^{(t)}\right)\left(\sigma_{t+1|t}\breve{z}_{i,1}\right)^2 + \sum_{j=1}^{\infty}\left(\frac{d^{(t)}\Gamma\left(j-d^{(t)}\right)}{\Gamma\left(1-d^{(t)}\right)\Gamma(j+1)}L^j\left(\varepsilon_{t+2|t+j}^2 - a_1^{(t)}\varepsilon_{t+1|t+j}^2\right)\right) + b_1^{(t)}\sigma_{t+1|t}^2 .$$

Note that $\sum_{j=1}^{\infty}\left(\frac{d^{(t)}\Gamma\left(j-d^{(t)}\right)}{\Gamma\left(1-d^{(t)}\right)\Gamma(j+1)}L^j\left(\varepsilon_{t+2|t+j}^2 - a_1^{(t)}\varepsilon_{t+1|t+j}^2\right)\right)$ is decomposed into

$$dL\left(\varepsilon_{t+2/t+1}^2 - a_1\varepsilon_{t+1/t+1}^2\right) + \frac{1}{2!}d(1-d)L^2\left(\varepsilon_{t+2/t+2}^2 - a_1\varepsilon_{t+1/t+2}^2\right) + \frac{1}{3!}d(1-d)(2-d)L^3\left(\varepsilon_{t+2/t+3}^2 - a_1\varepsilon_{t+1/t+3}^2\right) + \dots , \qquad \text{or}$$

$$d\left(\varepsilon_{t+1/t}^2 - a_1\varepsilon_{t/t}^2\right) + \frac{1}{2!}d(1-d)\left(\varepsilon_{t/t}^2 - a_1\varepsilon_{t-1/t}^2\right) + \frac{1}{3!}d(1-d)(2-d)\left(\varepsilon_{t-1/t}^2 - a_1\varepsilon_{t-2/t}^2\right) + \dots . \text{ Note also that the term}$$

$\varepsilon_{t+1/t}^2$, which is not estimable, is replaced by the simulated component $\left(\sigma_{t+1|t}\breve{z}_{i,1}\right)^2$. Therefore, the forecast variance for time $t+2$ can be computed as:

$$\breve{\sigma}_{i,t+2}^2 = a_0^{(t)} + \left(a_1^{(t)} - b_1^{(t)}\right)\left(\sigma_{t+1|t}\breve{z}_{i,1}\right)^2 + b_1^{(t)}\sigma_{t+1|t}^2 +$$

$$d\left(\left(\sigma_{t+1|t}\breve{z}_{i,1}\right)^2 - a_1\varepsilon_{t/t}^2\right) + \frac{1}{2!}d(1-d)\left(\varepsilon_{t/t}^2 - a_1\varepsilon_{t-1/t}^2\right) + \frac{1}{3!}d(1-d)(2-d)\left(\varepsilon_{t-1/t}^2 - a_1\varepsilon_{t-2/t}^2\right) + \dots .$$

Note 4: This is based on the 3-step ahead volatility forecasting formula:

$$\breve{\sigma}_{i,t+3}^2 = a_0^{(t)} + \left(a_1^{(t)} - b_1^{(t)}\right)\left(\sigma_{t+2|t}\breve{z}_{i,2}\right)^2 + \sum_{j=1}^{\infty}\left(\frac{d^{(t)}\Gamma\left(j-d^{(t)}\right)}{\Gamma\left(1-d^{(t)}\right)\Gamma(j+1)}L^j\left(\varepsilon_{t+3|t+j}^2 - a_1^{(t)}\varepsilon_{t+2|t+j}^2\right)\right) + b_1^{(t)}\sigma_{t+2|t}^2 .$$

Note that $\sum_{j=1}^{\infty}\left(\frac{d^{(t)}\Gamma\left(j-d^{(t)}\right)}{\Gamma\left(1-d^{(t)}\right)\Gamma(j+1)}L^j\left(\varepsilon_{t+3|t+j}^2 - a_1^{(t)}\varepsilon_{t+2|t+j}^2\right)\right)$ is decomposed into

$$dL\left(\varepsilon_{t+3/t+1}^2 - a_1\varepsilon_{t+2/t+1}^2\right) + \frac{1}{2!}d(1-d)L^2\left(\varepsilon_{t+3/t+2}^2 - a_1\varepsilon_{t+2/t+2}^2\right) + \frac{1}{3!}d(1-d)(2-d)L^3\left(\varepsilon_{t+3/t+3}^2 - a_1\varepsilon_{t+2/t+3}^2\right) + \dots , \qquad \text{or}$$

$$d\left(\varepsilon_{t+2/t}^2 - a_1\varepsilon_{t+1/t}^2\right) + \frac{1}{2!}d(1-d)\left(\varepsilon_{t+1/t}^2 - a_1\varepsilon_{t/t}^2\right) + \frac{1}{3!}d(1-d)(2-d)\left(\varepsilon_{t/t}^2 - a_1\varepsilon_{t-1/t}^2\right) + \dots . \text{ Note also that the}$$

terms $\varepsilon_{t+2/t}^2$ and $\varepsilon_{t+1/t}^2$, which are not estimable, are replaced by the simulated components $\left(\sigma_{t+2|t}\breve{z}_{i,2}\right)^2$ and $\left(\sigma_{t+1|t}\breve{z}_{i,1}\right)^2$, respectively. Therefore, the forecast variance for time $t+3$ can be computed as:

$$\breve{\sigma}_{i,t+3}^2 = a_0^{(t)} + \left(a_1^{(t)} - b_1^{(t)}\right)\left(\breve{\sigma}_{i,t+2}\breve{z}_{i,2}\right)^2 + b_1^{(t)}\breve{\sigma}_{i,t+2}^2 +$$

$$d\left(\left(\breve{\sigma}_{i,t+2}\breve{z}_{i,2}\right)^2 - a_1\left(\breve{\sigma}_{i,t+1}\breve{z}_{i,1}\right)^2\right) + \frac{1}{2!}d(1-d)\left(\left(\breve{\sigma}_{i,t+1}\breve{z}_{i,1}\right)^2 - a_1\varepsilon_{t/t}^2\right) + \frac{1}{3!}d(1-d)(2-d)\left(\varepsilon_{t/t}^2 - a_1\varepsilon_{t-1/t}^2\right) + \dots .$$

Table 1. Descriptive Statistics of the 10 indices.

INDEX	Observations	Mean	Median	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Probability[1]
ATXINDX	4945	0.019052	0.054244	1.408684	-0.28461	10.83047	12700.47	0.000
FRCAC40	5050	0.013500	0.033854	1.412313	-0.03769	7.751845	4752.411	0.000
DAXINDX	5044	0.025316	0.083406	1.48699	-0.12667	8.012514	5293.989	0.000
FTSE100	5051	0.016534	0.040528	1.150512	-0.11487	9.561969	9073.328	0.000
HNGKNGI	4944	0.042274	0.062797	1.725812	0.007238	12.05344	16884.78	0.000
JAPDOWA	4924	-0.02745	-0.01268	1.578394	-0.02021	8.277182	5713.949	0.000
NASA100	5031	0.041204	0.121181	1.900873	0.098471	7.94713	5138.512	0.000
NYSEALL	5035	0.025034	0.058148	1.122936	-0.3696	15.23759	31532.73	0.000
S&PCOMP	5039	0.022856	0.048103	1.17312	-0.19843	12.1539	17626.3	0.000
SWSEALI	4999	0.030183	0.066004	1.393286	0.148244	7.555802	4341.474	0.000

[1] This column displays the p-value for the Jarque-Bera test which has as its null hypothesis that the returns series follow a Gaussian distribution. The indices are: Austrian Traded Index (ATXINDX), French CAC 40 Index (FRCAC40), DAX 30 Performance (DAXINDX), UK Financial Times Stock Exchange 100 (FTSE100), Hang Seng (HNGKNGI), Nikkei 225 Stock Average (JAPDOWA), OMX Stockholm (SWSEALI), NASDAQ 100 (NASA100), NYSE Composite (NYSEALL) and S&P500 Composite (S&PCOMP).

Table 2. Full sample parameter estimates of the FIGARCH-skT model, for the 10 indices.

Coefficients	ATXINDX	FRCAC40	DAXINDX	FTSE100	HNGKNGI	JAPDOWA	NASA100	NYSEALL	S&PCOMP	SWSEALI
c_0	0.049571	0.04444	0.05938	0.039123	0.086038	0.00526	0.074679	0.048733	0.046164	0.082983
	[-3.173]	[3.017]	[4.158]	[3.541]	[4.686]	[0.3084]	[3.999]	[4.78]	[4.309]	[5.419]
c_1	0.113109	-0.00873	-0.01947	-0.01112	0.047015	-0.02945	-0.00292	0.011439	-0.02702	0.061998
	[7.58]	[-0.6305]	[-1.421]	[-0.7851]	[3.232]	[-2.15]	[-0.2095]	[0.8336]	[-2.012]	[4.128]
a_0	0.07989	0.049509	0.04517	0.034415	0.099726	0.066971	0.087961	0.025265	0.026566	0.050801
	[3.983]	[2.97]	[3.654]	[3.495]	[3.208]	[3.342]	[3.482]	[3.392]	[3.356]	[3.674]
d	0.406859	0.528927	0.567897	0.484713	0.403836	0.570475	0.441126	0.480434	0.492826	0.490655
	[7.978]	**[6.13]**	**[8.54]**	**[7.656]**	**[8.083]**	**[6.925]**	**[9.359]**	**[7.915]**	**[8.461]**	**[8.295]**
a_1	0.204813	0.146109	0.08897	0.138693	0.192626	0.096545	0.178037	0.165812	0.167933	0.180777
	[3.035]	[3.819]	[2.528]	[3.426]	[3.134]	[2.38]	[4.176]	[4.635]	[4.837]	[3.673]
b_1	0.476467	0.638456	0.625246	0.574383	0.531981	0.636711	0.587057	0.619576	0.642932	0.57433
	[5.823]	[8.39]	[10.38]	[7.915]	[6.311]	[8.479]	[9.565]	[9.882]	[11.15]	[8.536]
g	-0.06829	-0.07319	-0.10362	-0.07669	-0.06898	-0.04584	-0.07333	-0.07857	-0.06676	-0.04414
	[-3.304]	**[-3.488]**	**[-5.059]**	**[-3.489]**	**[-3.893]**	**[-2.344]**	**[-3.801]**	**[-4.372]**	**[-3.78]**	**[-2.029]**
v	7.083723	11.7244	8.959231	13.81621	6.887763	8.031668	12.27752	7.305493	7.478592	9.309369
	[11.04]	**[5.729]**	**[7.431]**	**[5.349]**	**[9.859]**	**[8.647]**	**[6.334]**	**[9.927]**	**[9.725]**	**[7.723]**

The numbers in brackets report the coefficient to standard error ratios. The indices are: Austrian Traded Index (ATXINDX), French CAC 40 Index (FRCAC40), DAX 30 Performance (DAXINDX), UK Financial Times Stock Exchange 100 (FTSE100), Hang Seng (HNGKNGI), Nikkei 225 Stock Average (JAPDOWA), OMX Stockholm (SWSEALI), NASDAQ 100 (NASA100), NYSE Composite (NYSEALL) and S&P500 Composite (S&PCOMP).

Table 3. 95% VaR and 95% ES forecasts for FIGARCH-skT model.

Index	Number of steps ahead	Number of VaR forecasts	Average VaR	Average ES	MSE	Observed exception rate	Independence Test p-value	Conditional Coverage Test p-value
ATXINDX	1	2945	-2.396	-3.309	0.028	3.50%	0.230	0.000**
	10	294	-2.109	-2.772	0.036	5.44%	0.266	0.509
	20	147	-2.044	-2.722	0.063	3.95%	0.699	0.592
FRCAC40	1	3050	-2.4964	-3.297	0.030	4.52%	0.276	0.261
	10	305	-2.5307	-3.2612	0.024	3.61%	0.400	0.353
	20	152	-2.4919	-3.277	0.071	3.95%	0.481	0.641
DAXINDX	1	3044	-2.797	-3.753	0.034	4.24%	0.815	0.137
	10	304	-2.877	-3.759	0.036	5.92%	0.080	0.167
	20	152	-2.773	-3.745	0.023	3.95%	0.521	0.671
FTSE100	1	3051	-2.059	-2.6962	0.01978	4.88%	0.313	0.575
	10	305	-1.969	-2.5609	0.06604	4.26%	0.572	0.709
	20	152	-1.9083	-2.5462	0.046	5.92%	0.089	0.208
HNGKNGI	1	2944	-3.051	-4.21	0.042	3.19%	0.570	0.000**
	10	294	-3.017	-3.905	0.018	4.76%	0.690	0.906
	20	147	-2.858	-3.766	0.027	5.44%	0.335	0.611
JAPDOWA	1	2924	-2.848	-3.854	0.036	3.42%	0.077	0.000**
	10	292	-2.928	-3.758	0.199	4.45%	0.598	0.789
	20	146	-2.804	-3.667	0.007	3.42%	0.550	0.544
NASA100	1	3031	-3.417	-4.51	0.045	4.26%	0.293	0.090
	10	303	-3.725	-4.834	0.071	4.62%	0.051	0.065

	20	151	-3.503	-4.641	0.047	4.64%	0.309	0.582
NYSEALL	1	3035	-2.152	-2.965	0.025	4.02%	0.965	0.037*
	10	303	-2.02	-2.62	0.042	5.28%	0.050	0.136
	20	151	-1.953	-2.59	0.031	5.96%	0.284	0.494
S&PCOMP	1	3039	-2.266	-3.106	0.031	3.85%	0.812	0.009**
	10	303	-2.096	-2.708	0.02	4.95%	0.210	0.456
	20	151	-2.032	-2.667	0.013	5.30%	0.421	0.715
SWSEALI	1	2999	-2.5	-3.337	0.027	4.37%	0.589	0.232
	10	299	-2.429	-3.192	0.043	5.02%	0.776	0.960
	20	149	-2.316	-3.12	0.071	5.37%	0.427	0.717

*denotes significant at 5%, ** denotes significant at 1%.

The indices are: Austrian Traded Index(ATXINDX), French CAC 40 Index (FRCAC40), DAX 30 Performance (DAXINDX), UKFinancial Times Stock Exchange 100 (FTSE100), Hang Seng (HNGKNGI), Nikkei 225Stock Average (JAPDOWA), OMX Stockholm (SWSEALI), NASDAQ 100 (NASA100), NYSEComposite (NYSEALL) and S&P500 Composite (S&PCOMP).

Table 4. 95% VaR and 95% ES forecasts for GARCH-skT model.

Index	Number of steps ahead	Number of VaR forecasts	Average VaR	Average ES	MSE	Observed exception rate	Independence Test p-value	Conditional Coverage Test p-value
ATXINDX	1	2945	-2.337	-3.229	0.031	3.60%	0.050	0.000**
	10	294	-1.943	-2.541	0.055	6.80%	0.042*	0.049*
	20	147	-1.907	-2.55	0.121	3.95%	0.151	0.228
FRCAC40	1	3050	-2.4621	-3.2441	0.031	4.92%	0.543	0.813
	10	305	-2.2403	-2.872	0.024	4.59%	0.666	0.862
	20	152	-2.2478	-2.9233	0.064	4.61%	0.409	0.692
DAXINDX	1	3044	-2.787	-3.739	0.032	4.47%	0.638	0.350
	10	304	-2.451	-3.186	0.055	8.88%	0.635	0.017*
	20	152	-2.449	-3.269	0.023	3.95%	0.521	0.671
FTSE100	1	3051	-2.0696	-2.7114	0.02025	4.85%	0.752	0.885
	10	305	-1.9022	-2.4489	0.06726	4.92%	0.762	0.953
	20	152	-1.9025	-2.4949	0.0513	5.26%	0.344	0.633
HNGKNGI	1	2944	-3.091	-4.281	0.038	3.13%	0.940	0.000**
	10	294	-2.707	-3.497	0.078	5.10%	0.207	0.450
	20	147	-2.724	-3.58	0.044	3.40%	0.552	0.536
JAPDOWA	1	2924	-2.837	-3.834	0.034	3.63%	0.935	0.002**
	10	292	-2.553	-3.268	0.224	5.82%	0.338	0.520
	20	146	-2.551	-3.314	0.01	4.11%	0.472	0.677
NASA100	1	3031	-3.497	-4.626	0.055	3.93%	0.876	0.018*
	10	303	-3.183	-4.075	0.085	6.27%	0.129	0.196

	20	151	-3.213	-4.175	0.071	5.30%	0.421	0.714
NYSEALL	1	3035	-2.164	-2.982	0.022	3.86%	0.799	0.009**
	10	303	-1.865	-2.406	0.072	5.94%	0.095	0.191
	20	151	-1.859	-2.445	0.031	6.62%	0.232	0.338
S&PCOMP	1	3039	-2.309	-3.171	0.026	3.72%	0.917	0.003**
	10	303	-2.009	-2.572	0.029	4.62%	0.243	0.481
	20	151	-2.036	-2.645	0.011	5.30%	0.421	0.715
SWSEALI	1	2999	-2.494	-3.337	0.03	4.63%	0.853	0.639
	10	299	-2.242	-2.941	0.057	6.02%	0.128	0.233
	20	149	-2.214	-2.976	0.103	4.70%	0.314	0.591

*denotes significant at 5%, ** denotes significant at 1%.

The indices are: Austrian Traded Index(ATXINDX), French CAC 40 Index (FRCAC40), DAX 30 Performance (DAXINDX), UKFinancial Times Stock Exchange 100 (FTSE100), Hang Seng (HNGKNGI), Nikkei 225Stock Average (JAPDOWA), OMX Stockholm (SWSEALI), NASDAQ 100 (NASA100), NYSEComposite (NYSEALL) and S&P500 Composite (S&PCOMP).

Table 5. The *p*-values of the SPA test for the null hypothesis that the FIGARCH-skT model outperforms the GARCH-skT model.

The loss function measures the squared distance between the actual and expected returns in the event of a 95% VaR violation

	Index									
Number of steps ahead	ATXINDX	FRCAC40	DAXINDX	FTSE100	HNGKNGI	JAPDOWA	NASA100	NYSEALL	S&PCOMP	SWESALI
1	0.82140	0.56110	0.16440	0.62630	0.14020	0.17460	0.95050	0.15310	0.06630	0.51180
10	0.55070	0.43820	0.53690	0.70940	0.57170	0.84780	0.52670	0.84910	0.80010	0.80980
20	0.55810	0.20700	0.58620	0.68520	0.75110	0.77190	0.55670	0.50480	0.16330	0.77420

The loss function measures the absolute distance between the actual and expected returns in the event of a 95% VaR violation

	Index									
Number of steps ahead	ATXINDX	FRCAC40	DAXINDX	FTSE100	HNGKNGI	JAPDOWA	NASA100	NYSEALL	S&PCOMP	SWESALI
1	0.61540	0.80200	0.40340	0.51600	0.35180	0.54330	0.60900	0.08270	0.12180	0.50620
10	0.54330	0.75590	0.52510	0.84540	0.54380	0.79920	0.52540	0.55330	0.52310	0.70530
20	0.53330	0.72660	0.39860	0.65350	0.50420	0.75220	0.56520	0.27420	0.23340	0.50020

The indices are: Austrian Traded Index(ATXINDX), French CAC 40 Index (FRCAC40), DAX 30 Performance (DAXINDX), UKFinancial Times Stock Exchange 100 (FTSE100), Hang Seng (HNGKNGI), Nikkei 225Stock Average (JAPDOWA), OMX Stockholm (SWSEALI), NASDAQ 100 (NASA100), NYSEComposite (NYSEALL) and S&P500 Composite (S&PCOMP).

Table 6. 99% VaR and 99% ES forecasts for FIGARCH-skT model.

Index	Number of steps ahead	Number of VaR forecasts	Average VaR	Average ES	MSE	Observed exception rate	Independence Test p-value	Conditional Coverage Test p-value
ATXINDX	1	2945	-3.840	-4.835	0.003	0.54%	0.676	0.022*
	10	294	-3.186	-3.818	0.006	2.04%	0.616	0.258
	20	147	-3.125	-3.777	0.026	3.40%	0.551	0.060
FRCAC40	1	3050	-3.785	-4.542	0.008	0.62%	0.108	0.021*
	10	305	-3.714	-4.343	0.008	0.66%	0.871	0.802
	20	152	-3.753	-4.472	0.019	1.97%	0.727	0.536
DAXINDX	1	3044	-4.329	-5.268	0.008	0.66%	0.121	0.039*
	10	304	-4.323	-5.149	0.016	1.32%	0.744	0.826
	20	152	-4.333	-5.298	0.001	1.97%	0.727	0.535
FTSE100	1	3051	-3.088	-3.671	0.004	1.08%	0.396	0.631
	10	305	-2.938	-3.481	0.039	1.64%	0.683	0.543
	20	152	-2.922	-3.521	0.014	2.63%	0.641	0.221
HNGKNGI	1	2944	-4.877	-6.163	0.009	0.54%	0.676	0.022*
	10	294	-4.465	-5.272	0.004	1.70%	0.677	0.503
	20	147	-4.322	-5.184	0.000	0.68%	0.907	0.911
JAPDOWA	1	2924	-4.445	-5.510	0.007	0.55%	0.675	0.024*
	10	292	-4.293	-5.029	0.126	2.05%	0.100	0.074
	20	146	-4.202	-5.029	0.001	1.37%	0.813	0.889
NASA100	1	3031	-5.177	-6.205	0.009	0.53%	0.073	0.003**
	10	303	-5.541	-6.520	0.015	0.66%	0.870	0.807

	20	151	-5.345	-6.414	0.011	1.99%	0.035*	0.062
NYSEALL	1	3035	-3.438	-4.317	0.003	0.53%	0.680	0.014*
	10	303	-2.998	-3.544	0.020	1.98%	0.622	0.283
	20	151	-2.982	-3.594	0.015	1.99%	0.726	0.532
S&PCOMP	1	3039	-3.597	-4.496	0.003	0.49%	0.700	0.007**
	10	303	-3.093	-3.635	0.000	0.66%	0.870	0.805
	20	151	-3.077	-3.674	0.001	1.32%	0.816	0.908
SWSEALI	1	2999	-3.842	-4.660	0.008	1.00%	0.436	0.738
	10	299	-3.663	-4.384	0.002	1.00%	0.805	0.970
	20	149	-3.614	-4.423	0.008	1.34%	0.815	0.901

*denotes significant at 5%, ** denotes significant at 1%.

The indices are: Austrian Traded Index(ATXINDX), French CAC 40 Index (FRCAC40), DAX 30 Performance (DAXINDX), UKFinancial Times Stock Exchange 100 (FTSE100), Hang Seng (HNGKNGI), Nikkei 225Stock Average (JAPDOWA), OMX Stockholm (SWSEALI), NASDAQ 100 (NASA100), NYSEComposite (NYSEALL) and S&P500 Composite (S&PCOMP).

Table 7. 99% VaR and 99% ES forecasts for GARCH-skT model.

Index	Number of steps ahead	Number of VaR forecasts	Average VaR	Average ES	MSE	Observed exception rate	Independence Test p-value	Conditional Coverage Test p-value
ATXINDX	1	2945	-3.747	-4.720	0.005	0.75%	0.565	0.299
	10	294	-2.917	-3.487	0.018	2.72%	0.503	0.040*
	20	147	-2.932	-3.557	0.059	4.08%	0.220	0.008**
FRCAC40	1	3050	-3.722	-4.455	0.009	0.62%	0.625	0.071
	10	305	-3.276	-3.823	0.008	0.66%	0.871	0.802
	20	152	-3.338	-3.940	0.016	1.97%	0.727	0.536
DAXINDX	1	3044	-4.313	-5.247	0.007	0.49%	0.062	0.001**
	10	304	-3.642	-4.317	0.024	1.64%	0.682	0.540
	20	152	-3.760	-4.536	0.001	2.63%	0.641	0.219
FTSE100	1	3051	-3.106	-3.694	0.004	1.08%	0.396	0.631
	10	305	-2.794	-3.287	0.041	1.64%	0.683	0.543
	20	152	-2.877	-3.428	0.010	2.63%	0.641	0.221
HNGKNGI	1	2944	-4.963	-6.296	0.006	0.51%	0.695	0.012*
	10	294	-3.989	-4.707	0.005	2.04%	0.616	0.257
	20	147	-4.115	-4.915	0.002	1.36%	0.814	0.892
JAPDOWA	1	2924	-4.421	-5.472	0.006	0.62%	0.637	0.071
	10	292	-3.716	-4.337	0.141	1.71%	0.676	0.496
	20	146	-3.776	-4.461	0.000	0.68%	0.906	0.914
NASA100	1	3031	-5.314	-6.387	0.013	0.43%	0.738	0.001**
	10	303	-4.640	-5.391	0.017	1.32%	0.743	0.822

	20	151	-4.736	-5.583	0.023	1.99%	0.726	0.531
NYSEALL	1	3035	-3.458	-4.347	0.002	0.53%	0.680	0.014*
	10	303	-2.753	-3.239	0.020	2.64%	0.509	0.047*
	20	151	-2.812	-3.368	0.012	0.66%	0.908	0.898
S&PCOMP	1	3039	-3.673	-4.603	0.003	0.53%	0.681	0.014*
	10	303	-2.938	-3.423	0.006	0.66%	0.870	0.805
	20	151	-3.025	-3.572	0.000	1.32%	0.816	0.908
SWSEALI	1	2999	-3.845	-4.676	0.008	0.87%	0.500	0.602
	10	299	-3.373	-4.024	0.015	1.67%	0.680	0.524
	20	149	-3.437	-4.203	0.030	2.01%	0.725	0.522

*denotes significant at 5%, ** denotes significant at 1%.

The indices are: Austrian Traded Index(ATXINDX), French CAC 40 Index (FRCAC40), DAX 30 Performance (DAXINDX), UKFinancial Times Stock Exchange 100 (FTSE100), Hang Seng (HNGKNGI), Nikkei 225Stock Average (JAPDOWA), OMX Stockholm (SWSEALI), NASDAQ 100 (NASA100), NYSEComposite (NYSEALL) and S&P500 Composite (S&PCOMP).

Table 8. 10-day ahead forecasts of 95% VaR and 95% ES for the FIGARCH-skT model under the square-root rule.

Index	Number of steps ahead	Number of VaR forecasts	Average VaR	Average ES	MSE	Observed exception rate	Independence Test p-value	Conditional Coverage Test p-value
				FIGARCH-skT model				
ATXINDX	10	294	-2.420	-3.338	0.024	3.40%	0.400	0.287
FRCAC40	10	305	-2.501	-3.304	0.022	3.61%	0.400	0.353
DAXINDX	10	304	-2.799	-3.757	0.061	5.26%	0.812	0.952
FTSE100	10	305	-2.055	-2.691	0.056	3.93%	0.483	0.528
HNGKNGI	10	294	-3.066	-4.228	0.068	3.74%	0.354	0.379
JAPDOWA	10	292	-2.858	-3.868	0.200	4.45%	0.598	0.789
NASA100	10	303	-3.406	-4.497	0.054	5.61%	0.010*	0.032*
NYSEALL	10	303	-2.153	-2.967	0.023	4.62%	0.670	0.869
S&PCOMP	10	303	-2.268	-3.109	0.010	2.64%	0.509	0.094
SWSEALI	10	299	-2.505	-3.344	0.018	4.01%	0.316	0.431

*denotes significant at 5%.

The indices are: Austrian Traded Index(ATXINDX), French CAC 40 Index (FRCAC40), DAX 30 Performance (DAXINDX), UKFinancial Times Stock Exchange 100 (FTSE100), Hang Seng (HNGKNGI), Nikkei 225Stock Average (JAPDOWA), OMX Stockholm (SWSEALI), NASDAQ 100 (NASA100), NYSEComposite (NYSEALL) and S&P500 Composite (S&PCOMP).

Figure 1a,b,c. Plots of observed returns, against the one-step, ten-step and twenty-step VaR forecasts produced by the FIGARCH and GARCH models for the SWSEALI index.
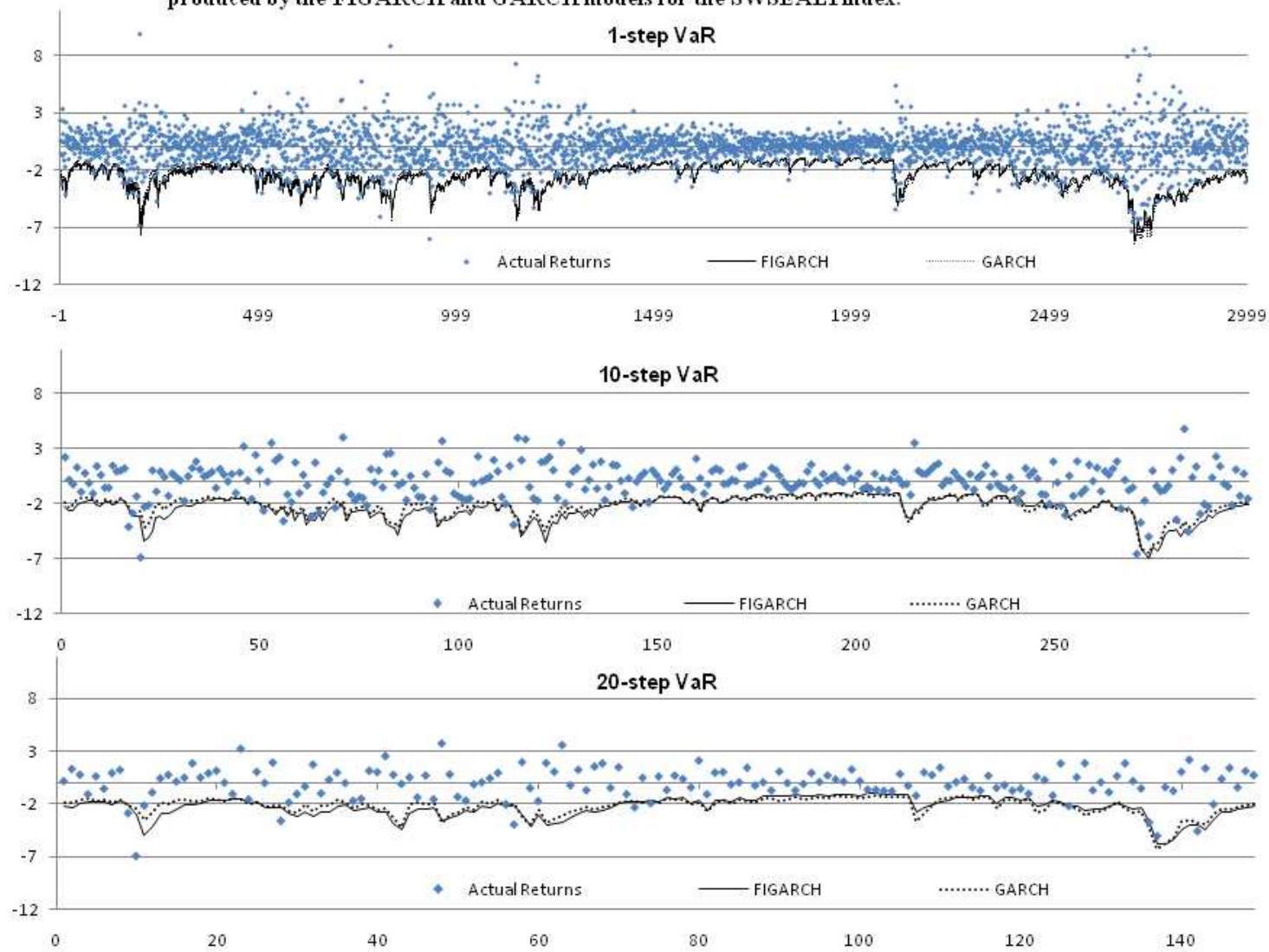
Figure 2a,b,c. Plots of observed returns, against the one-step, ten-step and twenty-step VaR forecasts produced by the FIGARCH and GARCH models for the HNGKNGI index