Value-at-Risk forecasting ability of filtered historical simulation for non-Normal GARCH returns

Chris Adcock (*) c.j.adcock@sheffield.ac.uk
Nelson Areal (**) nareal@eeg.uminho.pt
Benilde Oliveira (**) (***) benilde@eeg.uminho.pt

First Draft: February 2010
This Draft: January 2011

Abstract

Value-at-Risk (VaR) forecasting ability of Filtered Historical Simulation (FHS) is assessed using both simulated and empirical data. Three data generating processes are used to simulate several return samples. A backtesting is implemented to assess the performance of FHS based on a normal-GARCH. In addition, the performance of a GARCH model with *t*-Student and Skewed-*t* distributional assumptions for the residuals is also investigated. The simulation results are clearly in favour of the accuracy of FHS based on a normal-GARCH. Data on six well known active stock indices is used to produce empirical results. To evaluate FHS, four competing GARCH-type specifications, combined with three different innovation assumptions (normal, *t*-Stundent and Skewed-*t*), are used to capture time series dynamics. Though all the models demonstrate a good performance, the overall coverage results are in favour of the normal-GARCH. The use of GARCH models produces less favourable for FHS with respect to the independence of the VaR violations. The choice of an asymmetric GARCH structure to model the volatility dynamics of the empirical data results in a substantial improvement with respect to this issue. Furthermore, our results support the argument that distributionally nonparametric models do not depend on the distribution assumed in the filtering stage.

(***) Corresponding author

(*) University of Sheffield
Management School
Mappin Street
Sheffield, S1 4DT
United Kingdom
Tel: +44 (0)114 - 222 3402
Fax: +44 (0)114 - 222 3348

(**) NEGE – Management Research Unit
University of Minho
Campus de Gualtar, 4710-057 Braga,
Portugal
Tel: +351 253 604554
Fax: +351 253 601380

Value-at-Risk forecasting ability of filtered historical simulation for non-Normal GARCH returns

Abstract

Value-at-Risk (VaR) forecasting ability of Filtered Historical Simulation (FHS) is assessed using both simulated and empirical data. Three data generating processes are used to simulate several return samples. A backtesting is implemented to assess the performance of FHS based on a normal-GARCH. In addition, the performance of a GARCH model with *t*-Student and Skewed-*t* distributional assumptions for the residuals is also investigated. The simulation results are clearly in favour of the accuracy of FHS based on a normal-GARCH. Data on six well known active stock indices is used to produce empirical results. To evaluate FHS, four competing GARCH-type specifications, combined with three different innovation assumptions (normal, *t*-Stundent and Skewed-*t*), are used to capture time series dynamics. Though all the models demonstrate a good performance, the overall coverage results are in favour of the normal-GARCH. The use of GARCH models produces less favourable for FHS with respect to the independence of the VaR violations. The choice of an asymmetric GARCH structure to model the volatility dynamics of the empirical data results in a substantial improvement with respect to this issue. Furthermore, our results support the argument that distributionally nonparametric models do not depend on the distribution assumed in the filtering stage.

EFM classification: 370; 450; 530

1. Introduction

Value-at-Risk (VaR) synthesises in a single value the possible losses which could occur with a certain probability, in a given temporal horizon. Throughout the years, VaR became a standard downside measure of risk and has been receiving increasingly attention by academics and practitioners. Also, the accurate computation of VaR is critical for the estimation of other quantile-based risk measures such as the expected shortfall.

Traditionally the different methods to estimate VaR can be classified into two main categories: parametric methods (often denominated as local valuation methods) and simulation methods (Monte Carlo simulation and historical simulation). Typically, there is a trade-off between accuracy and the time required considering the application of these two different approaches. Parametric methods are less time consuming. However, when the time series under analysis exhibit important non-standard properties, simulation methods are more accurate. In fact, the adoption of full-valuation approaches to estimate VaR can lead to more accurate results as these methods generally depend on less restrictive distributional assumptions (see for example, Bucay and Rosen, 1999; Mausser and Rosen, 1999). As technical advances in relation to computational efficiency are likely to continue in the near future, the use of simulation methods is favoured.

Simplistic simulation methods based on the use of the empirical distribution to compute the tail quantiles, often referred to as historical simulation cannot adequately account for the volatility clustering. Therefore, these methods perform very poorly in practice. Recently, a new methodology has been developed in the literature to estimate VaR. This new method successfully combines bootstrapping techniques with the use of parametric models and is generally known as Filtered Historical Simulation (FSH). FHS was first proposed by Barone-Adesi et al. (1999). Under FHS the bootstrap process is applied to the residuals of a time series model used as a filter to extract autocorrelation and heteroscedasticity from the

historical time series of returns. Despite being numerically intensive, FHS is quite simple to apply and as a result it is faster to implement than other simulation methods. According to Hartz et al. (2006) FHS is also numerically extremely reliable. Additionally, FHS methodology makes no assumptions about the distribution of the returns under analysis. Based only on the assumption of IID standardised residuals from an appropriate volatility model, the use of the bootstrap algorithm allows a computationally feasible method to approximate the unknown return empirical distribution. This makes possible the computation of VaR with a great level of accuracy. In fact, Baroni-Adesi et al. (1999), Pritsker (2001) and most recently Kuester et al. (2005) have demonstrated the superiority of FHS method in the context of VaR estimation.

In recent research a new model based on heteroscedastic mixture distributions has been used in volatility modelling. This type of model links a GARCH-type structure to a discrete mixture of normal distributions, allowing for dynamic feedback between the components. The use of a mixture of normal's reduces the excess kurtosis so many times displayed by the residuals of traditional GARCH models. Haas et al. (2004) were pioneers in considering a mixed normal distribution associated with a GARCH-type structure (MN-GARCH). Later, Alexander and Lazar (2006) provided relevant evidence that generalized two-component MN-GARCH(1,1) models outperform those with three or more components, and also symmetric and skewed student's-t traditional GARCH models.

Kuester et al. (2005) compared the out-of-sample performance of existing alternative methods and some new models (including FHS and MN-GARCH) for predicting VaR in a univariate context. Using daily data for more than 30 years on the NASDAQ Composite Index, they found that most approaches perform inadequately. The only exceptions seem to be the GARCH-EVT (which focuses on tail estimation of the residuals of GARCH-filtered returns

via methods of extreme value theory), a new model based on heteroscedastic mixture distributions and the FHS.

Hartz et al. (2006) propose a new data driven method based on a classic normal GARCH model and use resampling methods to correct for the clear tendency of the model to underestimate the VaR. In fact, the suggested method is very flexible in accounting for the specific characteristics of the underlying series of returns as it is almost fully data driven. Their resampling method is related to the FHS method studied in Barone-Adesi et al. (1999, 2002). The results of Hartz et al. (2006) are also encouraging and demonstrate that the use of a simple normal GARCH model, combined with the application of a resampling method, may not need to be abandoned after all.

The remainder of the paper is organized as follows. Section 2 describes and discusses FHS as a method to estimate VaR. In section 3 we present and describe three different data generating processes (DGPs) used to simulate several time series samples. Additionally, the backtesting is described in detail. In section 4 the simulation and empirical results for the backtesting are reported. The final section, provides some concluding remarks.

2. Description of the Filtered Historical Simulation

FHS combines the best elements of conditional volatility models with the best elements of the simulation methods. By combining a non-parametric bootstrap procedure with a parametric modelling of the time series under analysis we must be able to considerable improve the overall quality of the VaR estimates.

The first step to the implementation of FHS consists in the adoption of a proper volatility model, usually a GARCH-type model with normally distributed residuals.

The major attraction of FHS is that it does not rely on any assumptions about the distribution of returns. As a non-parametric method FHS can accommodate skewness, excess kurtosis and other non-normal features presented by the empirical series of returns including a high percentage of zero-returns. In the context of FHS method the historical distribution of returns is used to estimate the portfolio´s VaR, assuming that it is a good proxy for the true distribution of returns over the next holding period (in our particular case, over the next day).

Let a sample of observed returns, $R_1, R_2, \dots, R_T$, be denoted by $\{R_t\}$. A bootstrap sample of size B, is denoted by $\{R_s^*\}$, where $R_s^* = R_{t_s}$ and t_s is an integer drawn from the set $1,2, \dots T$ using random sampling with replacement. From the bootstrap sample of returns the VaR can be easily estimated as the λ quantile from the bootstrap sample.

The bootstrap used in this study are similar to those carefully described in Barone-Adesi et al. (1999). The bootstrap can be easily applied to the (independent/uncorrelated) standardised residuals of any specification of the GARCH-type models. As an illustrative example a fully description of the bootstrap procedure based on a traditional normal-GARCH(1,1) model is presented next,

$$R_t = \mu + \varepsilon_t, with\ \varepsilon_t \sim N(0, \sigma_t^2), \tag{1}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{2}$$

Based on the estimation of the above model, the standardized residuals are defined as,

$$e_t = \frac{\hat{\varepsilon}_t}{\sqrt{\hat{\sigma}_t^2}}, \tag{3}$$

where $\hat{\epsilon}_t$ is the estimated residual and $\hat{\sigma}_t^2$ is the corresponding daily volatility estimate.

To start the simulation procedure we get the simulated innovation for the period $t+1$ (z_{t+1}^*) by drawing (with replacement) a random standardized residual (e_1^*) from the data set of historical residuals and scale it with the volatility of period $t+1$,

$$z_{t+1}^{*} = e_1^{*} \times \sqrt{\sigma_{t+1}^2}\,. \tag{4}$$

The variance of period $t+1$ can be estimated at the end of period t as,

$$\sigma_{t+1}^2 = \hat{\alpha}_0 + \hat{\alpha}_1 \varepsilon_t^2 + \hat{\beta}_1 \sigma_t^2, \tag{5}$$

where $\varepsilon_t$is the latest (for the last day in the sample) estimated residual and σ_t^2 is the latest (for the last day in the sample) estimated variance in equation (1) and (2) respectively.

The first simulated return for the period $t+1$ (R_{t+1}^{*}) can be computed as,

$$R_{t+1}^{*} = \hat{\mu} + z_{t+1}^{*}, \tag{6}$$

where z_{t+1}^{*} is the simulated residual for the period $t+1$.

This procedure can be repeated as many times as we want in order to generate a bootstrap sample, generally denoted by R^{*}, of any size size B. Finally, based on this bootstrap sample the filtered historical simulated VaR can be easily obtained as it corresponds to the λ quantile of the bootstrap sample generated under FHS.

3. Backtesting VaR forecasting ability of FHS for non-normal returns

In general, financial time series exhibit certain stylized patterns. These patterns are essentially fat tails, volatility clustering, long memory and asymmetry. Along the years the development of different models for volatility was guided by these stylized facts observed in the empirical data. The most attractive volatility models in applications, are probably the GARCH-type models.

A less standard feature that the financial data might exhibit is a high percentage of zero returns. Though less common, this particular feature might not be neglectful specially when we are dealing with daily data with respect to some specific markets. Paolella and Taschini (2008) conducted a study on the CO_2 emission allowance market and the daily return series

used by the author's exhibit a larger-than-usual number of zeros[1]. These authors remark that a high incidence of zeros in the empirical data precludes the use of traditional GARCH models to forecast VaR, even when they are applied under FHS. Though Paolella and Taschini (2008) recognize that FHS is, in general, a high effective method to estimate VaR, the authors claim that the forecasting performance of this method, critically relies on the adequacy of the innovation distributional assumption used for the GARCH filter applied to deal with heteroscedasticity in the data. The argument is that, if the residuals of the fitted GARCH model significantly departure from the assumed distributional form, FHS will fail to estimate with precision the empirical distribution of the returns. Therefore, any VaR computation based on FHS will be inaccurate.

The effective use of a GARCH model under a strictly analytical approach to compute VaR critically relies on the adequacy of the model distributional assumption. The high incidence of zeros in the return series, results in a data generating process that is not consistent with any typical distributional assumption. This means that, in the presence of a high incidence of zeros, the analytical solution for the VaR, based on the estimation of GARCH models, with a typical distributional assumption will be potentially biased. Paolella and Taschini (2008) claim that the very same argument applies even when a GARCH model is used only as a filter (to deal with heteroscedasticity in the series of returns) and a simulation solution is provided for the VaR instead of an analytical one.

Paolella and Taschini (2008) advocate the non-applicability of the FHS methodology because of the zeros-problem. As an alternative to accurately estimate VaR they propose a conditional analysis, whereby a mixture model is applied which properly accounts for both the GARCH-type effects and the zeros-problem.

[1] Paolella and Taschini (2008) report an incidence of 29% of zeros for their dataset.

We are not in agreement with the argument of Paolella and Taschini (2008). The main insight of the FHS method is that it is possible to capture the dynamics in the data (like for example conditional heteroscedasticity) and still be somewhat unrestrictive about the shape of the distribution of the returns. In fact, under FHS the distributional assumption with respect to the residuals is relaxed is favour of the much weaker assumption that the distribution is such that the parameters of the model can be consistently estimated. Therefore, in the contrary of Paolella and Taschini (2008), we argue that FHS is an accurate method to estimate VaR, even when the residuals of the GARCH filter clearly violate the distributional assumption due to the abundance of zeros.

3.1. Definition of alternative non-normal data generating processes

For testing the forecasting performance of FHS under controllable circumstances several time series of returns are simulated using three different data generating processes (DGP): a normal GARCH with a π proportion of zeros, a mixed-normal GARCH (MN-GARCH) and a Student's *t*-distributed asymmetric power ARCH (*t*-A-PARCH). By choosing these three different DGPs we aim to simulate series of non-normal returns that exhibit the usual stylized facts, such as conditional power tails for the residuals and asymmetric volatility responses (for the *t*-A-PARCH), sophisticated correlation dynamics and time-varying skewness (for the MN-GARCH) and also a significant incidence of zeros (for the GARCH combined with a high percentage of zeros).

3.1.1. GARCH model combined with a high percentage of zeros

We first consider a DGP under which a π proportion of zeros and a $1-\pi$ proportion of non-zero returns, is simulated. The non-zero returns are assumed to be uncorrelated[2] and to follow a traditional normal GARCH(1,1) process for the variance. Formally, the model that generates such a sample of simulated returns is given by,

$$R_t = \begin{cases} \mu + \varepsilon_t = \mu + \sigma_t z_t \; with \; z_t \sim N(0,1) \; and \; \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, for \; a \; proportion \; (1-\pi) \\ 0, for \; a \; proportion \; \pi \end{cases} \quad (7)$$

The proportion of zeros is set equal to 29% which corresponds to the incidence of zeros reported by Paolella and Taschini (2008).

In order to simulate the $1-\pi$ proportion of non-zero returns, typical values for the parameters are assumed in the model:

$$\mu = 0.0004,$$

$$\alpha_0 = 0.000006, \;\; \alpha_1 = 0.1, \;\; \beta_1 = 0.8,$$

3.1.2. MN-GARCH model

As an alternative, a MN-GARCH is used to generate the simulated returns. The functional form of a MN-GARCH time series R_t can be described as,

$$R_t = E(R_t|\Omega_{t-1}) + \varepsilon_t, \quad (8)$$

where $E(.|.)$ denotes the conditional expectation operator, $\Omega_{t-1}$the information set at time $t-1$, and ε_t the innovations or residuals of the time series. ε_t describes uncorrelated

[2] In interest of simplicity the returns are assumed to be uncorrelated and no ARMA terms will be included in the model.

disturbances with zero mean and plays the role of the unpredictable part of the return time series.

Under MN-GARCH models, the usual GARCH structure is extended by modelling the dynamics in volatility by a system of equations that enables feedback between the mixture components.

The time series of ε_t is said to be generated by a component MN-GARCH process if the conditional distribution of ε_t is an $n-$ component mixed normal with zero mean, that is,

$$\varepsilon_t|\Omega_{t-1} \sim MN(\omega, \mu, \sigma^2), \tag{9}$$

where ε_t denotes the innovations or residuals of a time series, Ω_{t-1} represents the information set available at time $t-1$ and $\omega = (\omega_1, \dots, \omega_n), \mu = (\mu_1, \dots, \mu_n), \sigma_t^2 = (\sigma_{1t}^2, \dots, \sigma_{nt}^2)$. The mixed normal density is given by,

$$f_{MN}(z; \omega, \mu, \sigma_t^2) = \sum_{j=1}^{n} \omega_j \phi\left(z; \mu_j, \sigma_{jt}^2\right), \tag{10}$$

where ϕ is the normal pdf, $\omega = (\omega_1, \dots, \omega_n)'$ is the set of component weights such that $\omega_j \in (0,1)$ and $\sum_{j=1}^{n} \omega_j = 1$, $\mu = (\mu_1, \dots, \mu_n)'$ is the set of component means such that, to ensure $E[\varepsilon_t] = 0, \mu_n = -\sum_{j=1}^{n-1}\left(\omega_j/\omega_n\right)\mu_j$, and $\sigma_t^2 = (\sigma_{1t}^2, \dots, \sigma_{nt}^2)' \in R_+^n$ are the positive component variances at time t.

The key aspect of a MN-GARCH model is that the n component variances σ_t^2 are allowed to evolve according to the GARCH-type structure.

One major advantage of using a MN-GARCH process lies in the fact that time-varying skewness is intrinsic in the model without requiring explicit specification of a conditional skewness process. MN-GARCH models are similar to Markov switching models but easier for use. In fact, the MN-GARCH model can be seen as the Markov switching GARCH model in a restricted form where the transition probabilities are independent of the past state.

In this paper a MN(2)-GARC(1,1) is used as a DGP. Typical values for the parameters of the model are used:

$$\mu = 0.050,$$

$$\alpha_{10} = 0.00005,\ \ \alpha_{11} = 0.00292,\ \ \beta_{11} = 0.9434,\ \ \omega_{11} = 0.6595$$

$$\alpha_{20} = 0.0002,\ \ \alpha_{21} = 0.0786,\ \ \beta_{21} = 0.9465,$$

3.1.3. *t*-A-PARCH model

Finally, a very competitive model for fitting asset returns is used to generate our series of simulated returns: the asymmetric power ARCH model (A-PARCH). This model was first introduced by Ding et al. (1993). It allows asymmetric volatility responses and also conditional power tails. Mittinik and Paolella (2000) recommend the combination of a A-PARCH model with the assumption of Student's *t*-distributed residuals to improve the competitiveness of the model for fitting asset returns.

A *t*-A-PARCH model is represented as:

$$R_t = \mu + \varepsilon_t, \varepsilon_t = z_t, z_t \sim t_\nu, \tag{11}$$

$$\sigma_t^\delta = \alpha_0 + \alpha_1(|\varepsilon_{t-1}| - \gamma\varepsilon_{t-1})^\delta + \beta_1\sigma_{t-1}^\delta. \tag{12}$$

The samples of simulated returns are generated using typical values for the parameters in the model:

$$\mu = 0.050,$$

$$\alpha_0 = 0.035,\ \ \alpha_1 = 0.200,\ \ \beta_1 = 0.700,$$

$$\gamma = -0.200,\ \ \delta = 1.600,\ \ \nu = 5.000.$$

3.2. Backtesting FHS

To assess the forecasting performance of FHS we follow Christofferson's (2003) framework. Additionally, a Dynamic Quantile (DQ) test as suggested by Engle and Manganelli (2004) is also applied.

By observing a series of ex ante VaR forecasts and ex post returns we can define a hit sequence of VaR violations as:

$$I_t = \begin{cases} 1, if\ R_t < -VaR_t \\ 0, if\ R_t > -VaR_t \end{cases} \quad (13)$$

If FHS is a correct forecasting model the hit sequence of violations should be:

$$H_0: I_t \sim iid \text{ Bernouli } (\lambda).$$

3.2.1. Christoffersen´s framework: test of unconditional coverage, independence and conditional coverage

To test null hypothesis define above we must first test if the fraction of violations obtained for our risk model $\hat{\lambda}$ is equal to the specified fraction λ (unconditional coverage). Under the unconditional coverage null hypothesis that $\hat{\lambda} = \lambda$, the likelihood of an iid Bernoulli (λ) is given by:

$$L(\lambda) = \prod_{t=1}^{T} (1-\lambda)^{1-I_{t+1}} \lambda^{I_{t+1}} = (1-\lambda)^{T_0} \lambda^{T_1}, \quad (14)$$

and the observed likelihood is given by:

$$L(\hat{\lambda}) = \prod_{t=1}^{T} (1-\hat{\lambda})^{1-I_{t+1}} \hat{\lambda}^{I_{t+1}} = (1-\hat{\lambda})^{T_0} \hat{\lambda}^{T_1}, \quad (15)$$

where T_0 and T_1 are the number of $0s$ and $1s$ in the sample. The observed fraction of violations in the sequence can be estimated as $\hat{\lambda} = T_1/T$.

The unconditional coverage hypothesis can be checked using a likelihood ratio (LR) test:

$$LR_{uc} = -2ln\left[L(\lambda)/L(\hat{\lambda})\right] \tag{16}$$

The test is asymptotically distributed as a $\mathcal{X}^2$ with one degree of freedom:

$$LR_{uc} = -2ln\left[L(\lambda)/L(\hat{\lambda})\right] \sim \mathcal{X}_1^2. \tag{17}$$

The P-value is computed as:

$$P_{uc} = 1 - F_{\mathcal{X}_1^2}(LR_{uc}),$$

where $F_{\mathcal{X}_1^2}(.)$ represents the cumulative density function of a $\mathcal{X}^2$ variable with one degree of freedom. Whenever the P-value is below the desired significance level the null hypothesis is rejected.

Christoffersen (2003) also establishes a test to the independence of the violations. For that purpose assume that the hit sequence is dependent over time and can be described as first-order Markov sequence with transition probability matrix:

$$\Lambda = \begin{bmatrix} 1-\lambda_{01} & \lambda_{01} \\ 1-\lambda_{11} & \lambda_{11} \end{bmatrix}$$

where $\lambda_{ij}$are the probabilities given by $\lambda_{ij} = \Pr(I_t = i \; and \; I_{t+1} = j)$, $i,j = 0,1$. For a sample of T observations, the likelihood function of the first-order Markov process is:

$$L(\Lambda) = (1-\lambda_{01})^{T_{00}}\lambda_{01}^{T_{01}}(1-\lambda_{11})^{T_{10}}\lambda_{11}^{T_{11}}, \tag{18}$$

where $T_{ij}, i,j = 0,1$ is yhe number of observations with a j following a i. The observed probabilities are given by:

$$\hat{\lambda}_{01} = \frac{T_{01}}{T_{00} + T_{01}}, \quad \hat{\lambda}_{11} = \frac{T_{11}}{T_{10} + T_{11}},$$

and

$$\hat{\lambda}_{00} = 1 - \hat{\lambda}_{01}, \quad \hat{\lambda}_{10} = 1 - \hat{\lambda}_{11}.$$

Under the null hypothesis the likelihood function is given by:

$$L(\hat{\Lambda}) = (1 - \hat{\lambda})^{T_0} \hat{\lambda}^{T_1}, \tag{19}$$

and the observed likelihood value is given by[3]:

$$L(\hat{\Lambda}) = (1 - \hat{\lambda}_{01})^{T_{00}} \hat{\lambda}_{01}^{T_{01}} (1 - \hat{\lambda}_{11})^{T_{10}} \hat{\lambda}_{11}^{T_{11}}. \tag{20}$$

We can test the independence hypothesis that $\lambda_{01} = \lambda_{11}$by applying a LR test:

$$LR_{ind} = -2ln[L(\hat{\lambda})/L(\hat{\Lambda})] \sim \mathcal{X}_1^2. \tag{21}$$

The P-value is given by:

$$P_{ind} = 1 - F_{\mathcal{X}_1^2}(LR_{ind}),$$

Finally, Christoffersen (2003) proposes a jointly test for independence and correct coverage (conditional coverage test):

$$LR_{cc} = -2ln[L(\lambda)/L(\hat{\Lambda})] = L_{uc} + L_{ind} \sim \mathcal{X}_2^2. \tag{22}$$

[3] For samples where $T_{11} = 0$, the likelihood function is given by:

$$L(\hat{\Lambda}) = (1 - \hat{\lambda}_{01})^{T_{00}} \hat{\lambda}_{01}^{T_{01}}.$$

The corresponding P-value is given by:

$$P_{cc} = 1 - F_{\mathcal{X}_2^2}(LR_{cc}).$$

3.2.2. Dynamic Quantile test

Christoffersen´s (2003) framework has some limitations because it only considers the first-order dependence when assessing the independence of the series of VaR violations. As a consequence, under this approach a series of VaR violations, that exhibits some higher-order dependences, might not be rejected because it does not have first-order dependence. Engle and Manganelli (2004) propose an alternative test based on a linear regression.

Define:

$$Hit_t \equiv I(R_t < -VaR_t). \tag{23}$$

To construct the test the following regression must be implemented:

$$Hit_t = \alpha_0 + \sum_{i=1}^{M} \eta_i Hit_{t-i} + \eta_{M+1} VaR_t + u_t, \tag{24}$$

where $\eta_i, i = 1, \dots, M+1$, are regression parameters and

$$u_t \begin{cases} -\lambda, & with\ probability\ 1-\lambda \\ 1-\lambda, & with\ probability\ \lambda. \end{cases}$$

Under the null hypothesis , $\lambda_0 = \lambda$ and $\eta_i = 0, i = 1, \dots, M+1$. Using the vector notation:

$$\mathrm{H} - \lambda 1 = \mathrm{X}\eta + u, \tag{25}$$

where $\eta_0 = \lambda_0 - \lambda$ and 1 denotes the vector of ones. Under the null hypothesis, the regressores in (24) should not have explanatory power, that is $H_0: \eta = 0$. Invoking an

appropriate central limit theorem the asymptotic distribution of the OLS estimator under the null hypothesis can be established:

$$\hat{\eta}_{OLS} = (\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'(\mathrm{H} - \lambda 1) \overset{a}{\sim} N\big(0, (\mathrm{X}'\mathrm{X})^{-1}\lambda - (1-\lambda)\big), \tag{26}$$

Engle and Manganelli (2004) derive the following Dynamic Quantile (DQ) test statistic:

$$\mathrm{DQ} = \frac{\hat{\eta}'_{OLS}\,\mathrm{X}'\mathrm{X}\,\hat{\eta}'_{OLS}}{\lambda(1-\lambda)} \overset{a}{\sim} \chi^2_{M+2.} \tag{27}$$

In our study, similarly to Kuester et al. (2005), two alternative specifications of the DQ test are used: $\mathrm{DQ}_{\mathrm{Hit}}$, under which a constant and four lagged hits are included in the regression; $\mathrm{DQ}_{\mathrm{VaR}}$, under which the contemporaneous VaR estimate is also included in the regression.

4. Backtesting results

The use of simulated data offers a valuable opportunity to evaluate the VaR forecasting ability of FHS under controllable, yet realistic, circumstances. However, as the validity of any method is best assessed using empirical data sets, in addition to the backtesting exercise based on simulated data, the performance of FHS is also evaluated using six different empirical time series.

4.1. Simulation results

Ten samples, of 3500 simulated returns, are generated using each of the three DGP described in section 3.1. The first 500 generated observations are leftover to avoid any starting-value problems. The estimation period is defined as $T = 1000$. With respect to each simulated series, $P = 2000$ out-of-sample $h = 1$ step-ahead VaR forecasts are computed using downfall probabilities $\lambda = 0.01, 0.02, \ldots, 0.1$ and $B = 20\,000$ bootstrap replications. Three competing models, differing in the innovations assumption, are used as a filter in the context

of FHS method: a normal-GARCH(1,1), a *t*-GARCH(1,1) and a *Skew-t*-GARCH(1,1)[4]. The main purpose is to investigate whether the performance of FHS based on a GARCH process is sensitive to the use of different distributional assumptions for the innovations. Barone-Adesi et al. (1999) describes FHS as a distributional free method. Therefore the VaR forecasting performance of FHS should not be affected by the decision on the distribution that is assumed in the filtering stage. However the conclusion of Kuester et al. (2005) is disturbing in the sense that they indicate that the choice of a Skewed-*t* assumption for the residuals of the GARCH model, used as a filter under FHS, is able to produce better-quality results. This issue deserves further investigation.

The results for the first of the 10 simulated series, based on the three different DGP are reported in table 1 to table 3. A summary for the number of rejections of the null hypothesis for all the simulated samples is reported in table 4 to table 6.

Based on the results reported, we can conclude that the FHS method is a very accurate method to forecast VaR. With respect to all DGP, for a significance level of 1%, the null hypothesis that the risk models are correct on average is generally not rejected. When the DQ test is applied the results are slightly worst. Also, for a significance level of 1%, the independence of the VaR violations is generally preserved. For a significance level of 5% and 10%, some of the samples exhibit problems in terms of independence of the VaR violations, especially when a *t*-GARCH or Skew-*t*-GARCH is used.

[4] The model is estimated using maximum likelihood (ML) approach according to the quasi-Newton method of Broyden, Fletcher, Goldfarb and Shanno. The results were obtained using Ox version 4.10 (see Doornik, 2007).

Table 1 Results for the simulated series based on a GARCH + 29% of zeros (sample 1)[a]:

Target downfall probability	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
					normal					
Observed downfall probability	**0.0110**	**0.0210**	**0.0320**	**0.0380**	**0.0475**	**0.0570**	**0.0705**	**0.0780**	**0.0870**	**0.0980**
P_{uc}	**0.6582**	**0.7513**	**0.6039**	**0.6454**	**0.6051**	**0.5691**	**0.9302**	**0.7407**	**0.6375**	**0.7649**
P_{ind}	**0.4744**	**0.1743**	**0.7968**	**0.7746**	**0.4424**	**0.6874**	**0.7018**	**0.5952**	**0.6013**	**0.6132**
P_{cc}	**0.7021**	**0.3780**	**0.8456**	**0.8634**	**0.6513**	**0.7842**	**0.9258**	**0.8221**	**0.7807**	**0.8416**
DQ_{Hit}	**0.5912**	**0.5366**	**0.9977**	**0.9785**	**0.7180**	**0.9161**	**0.9011**	**0.8964**	**0.8790**	**0.6244**
DQ_{VaR}	**0.5820**	**0.5104**	**0.9593**	**0.7234**	**0.5515**	**0.9126**	**0.9380**	**0.9225**	**0.9384**	**0.7039**
					t -Student					
Observed downfall probability	**0.0115**	**0.0210**	**0.0295**	**0.0385**	**0.0505**	**0.0585**	**0.0700**	**0.0790**	**0.0870**	**0.0955**
P_{uc}	**0.5102**	**0.7513**	**0.8954**	**0.7305**	**0.9184**	**0.7767**	**1.0000**	**0.8688**	**0.6375**	**0.4994**
P_{ind}	**0.4548**	**0.1743**	**0.0561**	**0.7787**	**0.6033**	**0.6216**	**0.4698**	**0.6569**	**0.6694**	**0.3946**
P_{cc}	**0.6089**	**0.3780**	**0.1599**	**0.9059**	**0.8691**	**0.8504**	**0.7701**	**0.8938**	**0.8169**	**0.5541**
DQ_{Hit}	**0.3419**	**0.7485**	**0.8558**	**0.9770**	**0.9316**	**0.8468**	**0.9120**	**0.9857**	**0.9385**	**0.9528**
DQ_{VaR}	**0.4145**	**0.7896**	**0.8185**	**0.6472**	**0.7392**	**0.8411**	**0.9420**	**0.9824**	**0.9732**	**0.9507**
					Skewed-*t*					
Observed downfall probability	**0.0110**	**0.0200**	**0.0305**	**0.0365**	**0.0505**	**0.0570**	**0.0685**	**0.0795**	**0.0855**	**0.0955**
P_{uc}	**0.6582**	**1.0000**	**0.8960**	**0.4177**	**0.9184**	**0.5691**	**0.7919**	**0.9342**	**0.4786**	**0.4994**
P_{ind}	**0.4744**	**0.1957**	**0.7875**	**0.7314**	**0.7448**	**0.6874**	**0.3532**	**0.4048**	**0.5547**	**0.5148**
P_{cc}	**0.7021**	**0.4330**	**0.9561**	**0.6789**	**0.9434**	**0.7842**	**0.6276**	**0.7044**	**0.6534**	**0.6439**
DQ_{Hit}	**0.5912**	**0.6213**	**0.9999**	**0.9124**	**0.8832**	**0.8362**	**0.9358**	**0.9420**	**0.9312**	**0.8941**
DQ_{VaR}	**0.5779**	**0.6095**	**0.9826**	**0.6025**	**0.7644**	**0.8547**	**0.9700**	**0.9652**	**0.9687**	**0.9452**

[a] Estimation period of $T = 1000$; $P = 2000$ out-of-sample $h = 1$ step-ahead VaR forecasts are computed using downfall probabilities $\lambda = 0.01, 0.02, \ldots, 0.1$ and $M = 20\,000$ bootstrap replications. Entries in the last 10 columns are the P-values of the respective tests. Bold type entries are not significant at the 1% level. For the computation of DQ_{Hit} the estimated regression includes a constant and four lagged violations. For the computation of DQ_{VaR} the contemporaneous estimate of VaR in also included in the regression. See section 3.2. for a description of the tests.

Table 2 Results for the simulated series based on a *t*-A-PARCH (sample 1) [a]:

Target downfall probability	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
					normal					
Observed downfall probability	**0.0110**	**0.0210**	**0.0315**	**0.0425**	**0.0495**	**0.0570**	**0.0680**	**0.0785**	**0.0935**	**0.1025**
P_{uc}	**0.6582**	**0.7513**	**0.6964**	**0.5721**	**0.9182**	**0.5691**	**0.7248**	**0.8042**	**0.5866**	**0.7104**
P_{ind}	**0.2373**	**0.2852**	**0.0585**	**0.0907**	**0.1623**	**0.0791**	**0.1898**	**0.3690**	**0.4318**	**0.6020**
P_{cc}	**0.4510**	**0.5373**	**0.1548**	**0.2038**	**0.3747**	**0.1820**	**0.3979**	**0.6477**	**0.6333**	**0.8147**
DQ_{Hit}	**0.2671**	**0.7515**	0.0086	**0.2043**	**0.1480**	**0.1311**	**0.1273**	**0.2979**	**0.3920**	**0.7460**
DQ_{VaR}	**0.0527**	**0.0495**	**0.0003**	**0.0008**	**0.0021**	**0.0036**	**0.0017**	**0.0042**	**0.0042**	**0.0107**
					t -Student					
Observed downfall probability	**0.0105**	**0.0215**	**0.0300**	**0.0410**	**0.0490**	**0.0585**	**0.0690**	**0.0795**	**0.0920**	**0.1010**
P_{uc}	**0.8236**	**0.6359**	**1.0000**	**0.8202**	**0.8369**	**0.7767**	**0.8606**	**0.9342**	**0.7554**	**0.8817**
P_{ind}	**0.2139**	**0.0137**	**0.0019**	0.0644	**0.0652**	**0.2092**	**0.6803**	**0.5646**	**0.4845**	**0.4362**
P_{cc}	**0.4506**	**0.0429**	0.0081	**0.1762**	**0.1788**	**0.4366**	**0.9046**	**0.8443**	**0.7461**	**0.7304**
DQ_{Hit}	**0.2364**	**0.0123**	0.0014	**0.0281**	**0.1076**	**0.5751**	**0.3178**	**0.2983**	**0.4589**	**0.7429**
DQ_{VaR}	**0.0676**	0.0008	0.0001	0.0031	**0.0288**	**0.1971**	**0.1197**	**0.1454**	**0.2125**	**0.3690**
					Skewed-*t*					
Observed downfall probability	**0.0115**	**0.0210**	**0.0300**	**0.0405**	**0.0495**	**0.0580**	**0.0690**	**0.0780**	**0.0915**	**0.0995**
P_{uc}	**0.5102**	**0.7513**	**1.0000**	**0.9093**	**0.9182**	**0.7050**	**0.8606**	**0.7407**	**0.8151**	**0.9405**
P_{ind}	**0.2617**	**0.0115**	0.0019	**0.0571**	**0.0731**	**0.1920**	**0.6803**	**0.4920**	**0.4604**	**0.4795**
P_{cc}	**0.4289**	**0.0390**	0.0081	**0.1626**	**0.1997**	**0.3975**	**0.9046**	**0.7477**	**0.7410**	**0.7767**
DQ_{Hit}	**0.2864**	0.0099	0.0014	**0.0597**	**0.1237**	**0.6380**	**0.4507**	**0.3713**	**0.4234**	**0.5550**
DQ_{VaR}	**0.0705**	0.0017	0.0002	**0.0123**	**0.0555**	**0.2471**	**0.1258**	**0.1079**	**0.1228**	**0.2268**

[a] See the note in table 1.

Table 3 Results for the simulated series based on a MN-GARCH (sample 1) [a]:

Target downfall probability	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
					normal					
Observed downfall probability	**0.0125**	**0.0215**	**0.0300**	**0.0355**	**0.0500**	**0.0575**	**0.0655**	**0.0735**	**0.0855**	**0.0955**
P_{uc}	**0.2794**	**0.6359**	**1.0000**	**0.2954**	**1.0000**	**0.6356**	**0.4256**	**0.2778**	**0.4786**	**0.4994**
P_{ind}	**0.3138**	**0.0137**	0.0095	**0.0488**	**0.3433**	**0.5138**	**0.1192**	**0.4227**	**0.5258**	**0.2213**
P_{cc}	**0.3355**	**0.0429**	**0.0345**	**0.0830**	**0.6383**	**0.7222**	**0.2163**	**0.4024**	**0.6362**	**0.3768**
DQ_{Hit}	**0.1876**	**0.0220**	**0.0255**	**0.2298**	**0.8432**	**0.8918**	**0.4154**	**0.5424**	**0.8634**	**0.6491**
DQ_{VaR}	0.0011	0.0029	0.0048	**0.0516**	**0.4231**	**0.6522**	**0.1823**	**0.2814**	**0.7870**	**0.7105**
					t -Student					
Observed downfall probability	**0.0110**	**0.0200**	**0.0285**	**0.0375**	**0.0450**	**0.0550**	**0.0655**	**0.0720**	**0.0840**	**0.0955**
P_{uc}	**0.6582**	**1.0000**	**0.6917**	**0.5643**	**0.2970**	**0.3400**	**0.4256**	**0.1803**	**0.3435**	**0.4994**
P_{ind}	**0.0220**	**0.0506**	**0.0258**	**0.0792**	**0.3179**	**0.7360**	**0.1192**	**0.3502**	**0.5943**	**0.2213**
P_{cc}	**0.0658**	**0.1479**	**0.0770**	**0.1814**	**0.3525**	**0.5992**	**0.2163**	**0.2635**	**0.5541**	**0.3768**
DQ_{Hit}	0.0026	**0.1492**	**0.0950**	**0.4856**	**0.5387**	**0.7415**	**0.5278**	**0.5192**	**0.7433**	**0.4131**
DQ_{VaR}	0.0002	**0.0260**	**0.0237**	**0.2049**	**0.2447**	**0.5193**	**0.4651**	**0.4740**	**0.5796**	**0.5143**
					Skewed-*t*					
Observed downfall probability	**0.0105**	**0.0210**	**0.0300**	**0.0370**	**0.0470**	**0.0555**	**0.0660**	**0.0725**	**0.0870**	**0.0950**
P_{uc}	**0.8236**	**0.7513**	**1.0000**	**0.4882**	**0.5342**	**0.3911**	**0.4793**	**0.2096**	**0.6375**	**0.4527**
P_{ind}	**0.0180**	**0.0662**	**0.0395**	**0.0705**	**0.4162**	**0.7300**	**0.1315**	**0.3738**	**0.4535**	**0.2385**
P_{cc}	**0.0594**	**0.1760**	**0.1202**	**0.1532**	**0.5923**	**0.6522**	**0.2497**	**0.3065**	**0.6757**	**0.3765**
DQ_{Hit}	0.0016	**0.2062**	**0.2307**	**0.4288**	**0.7193**	**0.7822**	**0.5657**	**0.5005**	**0.8338**	**0.4096**
DQ_{VaR}	0.0003	**0.0286**	**0.0560**	**0.1533**	**0.2451**	**0.3456**	**0.3436**	**0.2904**	**0.5502**	**0.4703**

[a] See the note in table 1.

Table 4 Results for the simulated series based on a GARCH + 29% of zeros:
Number of rejections of the null hypothesis (number of P-values below the desired level)

	Significance level: 0.1					Significance level: 0.05					Significance level: 0.01				
Sample	P_{uc}	P_{ind}	P_{cc}	DQ_{Hit}	DQ_{VaR}	P_{uc}	P_{ind}	P_{cc}	DQ_{Hit}	DQ_{VaR}	P_{uc}	P_{ind}	P_{cc}	DQ_{Hit}	DQ_{VaR}
							normal								
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
3	0	2	2	1	1	0	2	0	0	1	0	0	0	0	0
4	0	2	1	2	2	0	1	0	1	1	0	0	0	1	1
5	0	2	1	0	1	0	1	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	2	0	2	2	0	0	0	2	1	0	0	0	1	0
8	0	0	0	3	2	0	0	0	2	1	0	0	0	1	1
9	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
							t -Student								
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	2	2	1	1	0	2	0	0	1	0	0	0	0	0
4	0	2	1	2	2	0	1	0	2	2	0	0	0	2	2
5	0	2	1	0	2	0	1	0	0	1	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	1	3	3	5	5	0	3	3	5	4	0	0	0	2	0
8	0	0	0	3	3	0	0	0	3	3	0	0	0	3	3
9	0	0	0	3	3	0	0	0	2	2	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
							Skewed-*t*								
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	2	2	1	1	0	2	0	0	0	0	0	0	0	0
4	0	2	1	2	2	0	1	0	2	2	0	0	0	2	2
5	0	2	1	0	1	0	1	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	2	2	4	4	0	2	2	4	4	0	0	0	1	0
8	0	0	0	3	3	0	0	0	3	3	0	0	0	3	3
9	0	0	0	3	2	0	0	0	1	1	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 5 Results for the simulated series based on a *t*-A-PARCH:
Number of rejections of the null hypothesis (number of P-values below the desired level)

	Significance level: 0.1					Significance level: 0.05					Significance level: 0.01				
Sample	P_{uc}	P_{ind}	P_{cc}	DQ_{Hit}	DQ_{VaR}	P_{uc}	P_{ind}	P_{cc}	DQ_{Hit}	DQ_{VaR}	P_{uc}	P_{ind}	P_{cc}	DQ_{Hit}	DQ_{VaR}
							normal								
1	0	3	0	1	10	0	0	0	1	9	0	0	0	1	7
2	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
3	0	5	4	5	5	0	4	1	5	5	0	0	0	0	5
4	0	5	5	0	0	0	5	3	0	0	0	2	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	9	5	4	8	0	5	2	4	6	0	2	0	3	3
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	2	4	0	0	0	2	2	0	0	0	0	0
9	0	2	0	0	2	0	0	0	0	0	0	0	0	0	0
10	0	0	0	2	2	0	0	0	2	2	0	0	0	0	0
							t -Student								
1	0	4	2	3	5	0	2	2	3	4	0	1	1	1	3
2	1	1	1	1	1	0	1	0	1	1	0	0	0	1	1
3	0	2	1	2	6	0	2	0	2	4	0	0	0	0	2
4	0	6	5	1	1	0	5	3	0	0	0	1	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	8	7	5	6	0	6	5	4	5	0	2	1	1	4
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	6	1	1	1	0	2	0	0	0	0	0	0	0	0
9	0	2	1	1	1	0	1	0	0	0	0	0	0	0	0
10	0	1	0	3	2	0	0	0	1	1	0	0	0	0	0
							Skewed-*t*								
1	0	4	2	3	5	0	2	2	2	3	0	1	1	2	2
2	1	1	1	1	1	0	1	0	1	1	0	0	0	1	1
3	0	3	1	3	6	0	2	0	2	6	0	0	0	1	2
4	0	5	5	0	0	0	5	1	0	0	0	1	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	7	6	5	6	0	7	5	5	5	0	4	2	2	4
7	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
8	0	5	3	2	2	0	3	2	2	2	0	1	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	2	2	0	0	0	2	2	0	0	0	0	0

Table 6 Results for the simulated series based on a MN-GARCH:
Number of rejections of the null hypothesis (number of P-values below the desired level)

	Significance level: 0.1					Significance level: 0.05					Significance level: 0.01				
Sample	P_{uc}	P_{ind}	P_{cc}	DQ_{Hit}	DQ_{VaR}	P_{uc}	P_{ind}	P_{cc}	DQ_{Hit}	DQ_{VaR}	P_{uc}	P_{ind}	P_{cc}	DQ_{Hit}	DQ_{VaR}
							normal								
1	0	3	3	2	4	0	3	2	2	3	0	1	0	0	3
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	9	9	0	1	0	9	8	0	0	0	7	1	0	0
4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	3	1	3	1	0	2	1	1	1	0	0	0	0	0
6	0	6	6	1	7	0	6	6	0	5	0	5	2	0	2
7	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	2	2	4	4	0	2	2	4	4	0	1	0	0	1
							t -Student								
1	0	4	2	2	3	0	2	0	1	3	0	0	0	1	1
2	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
3	0	9	9	0	0	0	9	8	0	0	0	6	1	0	0
4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	3	1	2	1	0	1	1	1	1	0	0	0	1	1
6	0	6	6	1	5	0	6	6	0	4	0	5	2	0	1
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	5	0	0	0	0	2	0	0	0	0	1
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	2	2	3	4	0	2	0	1	1	0	0	0	0	0
							Skewed-*t*								
1	0	4	1	1	3	0	2	0	1	2	0	0	0	1	1
2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	9	9	0	0	0	9	8	0	0	0	6	1	0	0
4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	3	1	2	2	0	1	1	1	1	0	0	0	1	1
6	0	8	7	0	5	0	7	7	0	5	0	6	2	0	1
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	7	0	0	0	0	4	0	0	0	0	1
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	2	2	4	4	0	2	0	1	2	0	0	0	0	0

Overall, for a significance level of 1%, the FHS filtered by the traditional normal-GARCH gives the best results for the simulated time series generated by using a GARCH+29% of zeros as the DGP. For the data that was simulated using a DGP based on a *t*-A-PARCH, for a significance level of 1%, the traditional normal-GARCH still produces the best results in the context of the Christoffersen´s (2003) framework. Neverthless, when DQ tests are applied the FHS method coupled with a Skew-*t*-GARCH performs slightly better. Finally, for the series of simulated return based on a MN-GARCH, for a 1% level of significance, the use of a *t*-GARCH(1,1) or a Skew-*t*-GARCH, to filter the returns under FHS, slightly improves the results in terms of independence of the VaR violations. However, the results are not conclusive as the DQ_{Hit} test indicates that the traditional normal-GARCH is preferred.

FHS using a traditional normal-GARCH model predicts with great accuracy the downfall risk for the time series generated by a normal-GARCH with a 29% of zeros. These results are robust to the different statistic tests that were applied. In fact, for all the tests and across all the significance levels, the results are in favour the accuracy of FHS as a method to forecast VaR when the time series exhibit a high incidence of zeros. Remember that Paolella and Taschini (2008) argued the opposite.

Even when more sophisticated DGPs are used to simulate the data, the forecasting ability of FHS based on the traditional normal-GARCH model is still very satisfactory, especially for a significance level of 1%. The less satisfactory results respect to the independence of the VaR violations (especially for a significance level of 5% and 10%) for the series generated by a *t*-A-PARCH and MN-GARCH DGP. The use of a *t*-GARCH or Skew-*t*-GARCH instead of the traditional normal-GARCH does not generally (considering the three DGP´s) improves the results of FHS method.

4.2 Empirical results

The performance of FHS is also evaluated using empirical time series. For that purpose, the backtesting was repeated using six well known active stock indices from developed markets: DAX, DJI, FTSE, NASDAQ Composite, Nikkei 225 and S&P 500. The data set[5] consists of daily closing

[5] The data was obtained from DataStream.

levels of the indices from February the 8th , 1971 to September the 30th , 2010, yielding a total of 10 343 observations of daily log returns. Table 7 summarizes the descriptive statistics for the six stock indices investigated. All the time series exhibit stylized characteristics that indicate a potential departure from normality: negatively skewness and excess kurtosis.

Table 7 Summary of descriptive statistics for the stock indices

	DAX	DJI	FTSE	NASDAQ	Nikkei	S&P500
Sample Size	10 343	10 343	10 343	10 343	10 343	10 343
Minimum	-0.137099	-0.256320	-0.121173	-0.120478	-0.161354	-0.228330
Maximum	0.107975	0.105083	0.089434	0.132546	0.132346	0.109572
Mean	0.000242	0.000242	0.000296	0.000305	0.000143	0.000238
Standard Deviation	0.012513	0.010726	0.010753	0.012385	0.012585	0.010755
Skewness	-0.306679	-1.356437	-0.278622	-0.281302	-0.350004	-1.077088
Kurtosis	8.031591	38.387919	9.141047	10.549319	10.796472	27.641346

Considering the empirical properties of the data, in accordance to what is generally recommended by the literature, an autoregressive process of first-order (AR(1)) is used to model the stock indices returns:

$$R_t = \mu + \varphi R_{t-1} + \varepsilon_t. \tag{28}$$

To capture volatility dynamics, the innovations of the above model are modelled in the context of a GARCH-type process. Four alternative GARCH-type structures are considered. In addition to the traditional GARCH, a GARCH-in-mean (GARCH-M), the Glosten, Jajannathan and Runkle (GJR) and a GJR-in-mean (GJR-M) specifications for the variance are also tested.

The GARCH-M model (Engle et al., 1987) extends the traditional GARCH framework by allowing the conditional mean stock return to depend on its own conditional variance:

$$\mathrm{R_t} = \mu + \varphi \mathrm{R_{t-1}} + \varpi\sigma^2 + \varepsilon_t, \tag{29}$$

An important reason in financial theory for using a GARCH-M structure is because it deals to a certain degree, with the presence of conditional left skewness commonly observed in stock returns time series. This is related to the volatility feedback effect, which amplifies the impact of bad news but

dampens the impact of good news[6]. Volatility feedback therefore implies that stock price movements are correlated with volatility and the GARCH-M model incorporates that its specification.

The traditional GARCH model is unable to capture volatility asymmetry, which also usually characterises stock markets. In this regard the GJR model, as introduced by Glosten et al. (1993), is also considered. The GJR model allows for possible asymmetric behaviour in volatility. An indicator function is incorporated in the volatility specification of the model, enabling the volatility to react differently depending on the sign of past innovations:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \vartheta \varepsilon_{t-1}^2 I_{t-1}, \quad (30)$$

where $I_{t-1} = \begin{cases} 1 \text{ if } \varepsilon_{t-1} < 0 \\ 0 \text{ if } \varepsilon_{t-1} \geq 0. \end{cases}$

In order to estimate the models, it is necessary to make assumptions about the distribution of the error term. For each model, and similarly to what was done before in the context of the simulation, three error distributional assumptions are investigated for each of the four models: normal, *t*-Student and Skewed-*t*. Therefore, twelve different GARCH-type models, with an AR(1) term in the mean equation, are tested:

- normal-AR(1)-GARCH(1,1), henceforth normal-GARCH ;
- *t*-AR(1)-GARCH(1,1), henceforth *t*-GARCH;
- Skew-*t*-AR(1)-GARCH(1,1), henceforth Skew-*t*-GARCH;
- normal-AR(1)-GARCH(1,1)-M, henceforth normal-GARCH-M;
- *t*-AR(1)-GARCH(1,1)-M, henceforth *t*-GARCH-M;
- Skew-*t*-AR(1)-GARCH(1,1)-M, henceforth Skew-*t*-GARCH-M;
- normal-AR(1)-GJR(1,1), henceforth normal-GJR;
- *t*-AR(1)-GJR(1,1), henceforth *t*-GJR;
- Skew-*t*-AR(1)-GJR(1,1), henceforth Skew-*t*-GJR;

[6] For a detailed explanation of the volatility feedback effect see Campbell and Hentschel (1992).

- normal-AR(1)-GJR(1,1)-M, henceforth normal-GJR-M;
- *t*-AR(1)-GJR(1,1)-M, henceforth *t*-GJR-M;
- Skew-*t*-AR(1)-GJR(1,1)-M, henceforth Skew-*t*-GJR-M;

A moving window of 1000 observation is used and the model parameters are updated for each moving window based on one day increments. For each empirical time series, 9343 out-of-sample $h = 1$ step-ahead VaR forecasts are computed using target downfall probabilities of $\lambda = 0.01,0.02,\ldots,0.1$ and $B = 20\,000$ bootstrap replications. The results are reported in Tables 8, 9, 10 and 11.

According to the three-zone approach defined by the Basle Committee (1996), a VaR model is considered accurate (green zone) if it produces a number of 1% VaR violations that remains below the binomial (0.01) 95% quantile. A model is arguable (yellow zone) up to 99.99% quantile. When more violations occur the model is judged as inappropriate (red zone). Adapting this framework to our sample size, if at most 109 (1.17%) violations occur the model is considered acceptable. Between 110 and 130 (1.39%) the model is classified as being disputable. Table 8 reports the observed percentage of VaR violations for all the models and across all target downfall probabilities. From the total of 72 estimated models (12 models for 6 different stock indices) 52 (approximately 72.2%) are classified as accurate (green), 19 (approximately 26.4%) as arguable (yellow) and only 1 (approximately 1.4%) is reported as inappropriate.

To expedite the comparison of VaR forecasting methods, Kuester et al. (2005) advocate the use of a graphical depiction of the quality of VaR predictions over the relevant probability levels. Therefore the relative deviation from the correct coverage can be compared across the different VaR levels and alternative models. The graphical depiction of the quality of the VaR forecasts was

implemented for all the stock indices investigated[7]. As an example, Figure 1 depicts the coverage results for VaR levels $\lambda = 0.01, 0.02, \ldots, 0.1$, with respect to the NASDAQ sample.

Based on the analysis of the deviation plots constructed for the all the empirical samples we can conclude that FHS has good coverage properties, , especially at the lower quantile. These result is valid across all the competing GARCH-type models. Nevertheless, in general, the traditional normal-GARCH model demonstrates a higher performance which indicates that this model is suitable to filter the empirical data for heteroscedasticity under FHS. The adoption of a more sophisticated GARCH-type model or distributional innovation assumption, does not generally improves the coverage properties of the FHS method to estimate VaR.

Table 9 gives summary information about to the coverage properties for each GARCH-type model investigated, across all the empirical samples. The mean absolute error (MAE) and the mean squared error (MSE) of the actual violation frequencies from the corresponding theoretical VaR level are reported.

[7] To save space the graphical depiction of the quality of VaR predictions, for all the stock indices samples, was not included in the paper. These results are available upon request.

Table 8 Percentage of VaR violations

Target downfall probability	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**	**0.1**	*
					DAX						
normal-GARCH	0.0110	0.0223	0.0334	0.0417	0.0519	0.0636	0.0730	0.0821	0.0907	0.1017	**Green**
t-GARCH	0.0115	0.0223	0.0334	0.0421	0.0535	0.0626	0.0734	0.0824	0.0914	0.1011	**Green**
Skew-*t*-GARCH	0.0116	0.0225	0.0333	0.0422	0.0530	0.0634	0.0727	0.0827	0.0922	0.1014	**Green**
normal-GARCH-M	0.0100	0.0207	0.0303	0.0381	0.0478	0.0582	0.0667	0.0751	0.0835	0.0916	**Green**
t-GARCH-M	0.0104	0.0206	0.0307	0.0390	0.0488	0.0580	0.0682	0.0759	0.0852	0.0933	**Green**
Skew-*t*-GARCH-M	0.0106	0.0208	0.0320	0.0389	0.0486	0.0584	0.0683	0.0760	0.0855	0.0942	**Green**
normal-GJR	0.0118	0.0223	0.0335	0.0423	0.0548	0.0637	0.0720	0.0807	0.0909	0.1022	**Yellow**
t-GJR	0.0124	0.0224	0.0336	0.0438	0.0548	0.0635	0.0728	0.0821	0.0931	0.1037	**Yellow**
Skew-*t*-GJR	0.0119	0.0223	0.0338	0.0433	0.0546	0.0638	0.0724	0.0822	0.0932	0.1034	**Yellow**
normal-GJR-M	0.0119	0.0225	0.0317	0.0408	0.0514	0.0610	0.0684	0.0770	0.0871	0.0971	**Yellow**
t-GJR-M	0.0109	0.0208	0.0311	0.0404	0.0492	0.0585	0.0669	0.0759	0.0862	0.0963	**Green**
Skew-*t*-GJR-M	0.0111	0.0210	0.0316	0.0407	0.0503	0.0590	0.0676	0.0771	0.0867	0.0971	**Green**
					DJI						
normal-GARCH	0.0118	0.0226	0.0332	0.0426	0.0535	0.0627	0.0715	0.0812	0.0896	0.0991	**Yellow**
t-GARCH	0.0111	0.0224	0.0324	0.0423	0.0515	0.0609	0.0710	0.0818	0.0900	0.1009	**Green**
Skew-*t*-GARCH	0.0119	0.0226	0.0328	0.0424	0.0515	0.0610	0.0716	0.0815	0.0900	0.1002	**Yellow**
normal-GARCH-M	0.0112	0.0210	0.0292	0.0374	0.0475	0.0547	0.0621	0.0710	0.0804	0.0881	**Green**
t-GARCH-M	0.0108	0.0216	0.0300	0.0382	0.0478	0.0566	0.0661	0.0755	0.0836	0.0927	**Green**
Skew-*t*-GARCH-M	0.0109	0.0217	0.0301	0.0394	0.0478	0.0560	0.0658	0.0760	0.0843	0.0927	**Green**
normal-GJR	0.0122	0.0233	0.0341	0.0442	0.0521	0.0639	0.0737	0.0824	0.0919	0.1001	**Yellow**
t-GJR	0.0112	0.0235	0.0330	0.0429	0.0520	0.0613	0.0724	0.0826	0.0917	0.0991	**Green**
Skew-*t*-GJR	0.0113	0.0239	0.0329	0.0430	0.0524	0.0614	0.0721	0.0825	0.0912	0.0996	**Green**
normal-GJR-M	0.0122	0.0222	0.0310	0.0408	0.0483	0.0572	0.0654	0.0736	0.0835	0.0920	**Yellow**
t-GJR-M	0.0113	0.0224	0.0313	0.0408	0.0490	0.0564	0.0658	0.0763	0.0843	0.0934	**Green**
Skew-*t*-GJR-M	0.0110	0.0220	0.0322	0.0408	0.0490	0.0573	0.0665	0.0771	0.0850	0.0931	**Green**
					FTSE						
normal-GARCH	0.0108	0.0198	0.0296	0.0407	0.0507	0.0598	0.0701	0.0802	0.0908	0.1004	**Green**
t-GARCH	0.0112	0.0203	0.0292	0.0402	0.0504	0.0606	0.0703	0.0804	0.0899	0.0999	**Green**
Skew-*t*-GARCH	0.0112	0.0204	0.0295	0.0398	0.0505	0.0595	0.0701	0.0802	0.0897	0.0995	**Green**
normal-GARCH-M	0.0110	0.0196	0.0274	0.0383	0.0485	0.0573	0.0676	0.0765	0.0849	0.0957	**Green**
t-GARCH-M	0.0105	0.0199	0.0286	0.0385	0.0485	0.0577	0.0683	0.0771	0.0853	0.0957	**Green**
Skew-*t*-GARCH-M	0.0104	0.0201	0.0284	0.0377	0.0487	0.0572	0.0686	0.0776	0.0852	0.0952	**Green**
normal-GJR	0.0113	0.0209	0.0304	0.0411	0.0519	0.0618	0.0710	0.0796	0.0912	0.1005	**Green**
t-GJR	0.0112	0.0213	0.0307	0.0418	0.0517	0.0626	0.0715	0.0811	0.0911	0.1008	**Green**
Skew-*t*-GJR	0.0112	0.0216	0.0307	0.0405	0.0516	0.0627	0.0713	0.0803	0.0913	0.1006	**Green**
normal-GJR-M	0.0111	0.0204	0.0294	0.0387	0.0507	0.0602	0.0696	0.0783	0.0880	0.0974	**Green**
t-GJR-M	0.0108	0.0203	0.0295	0.0390	0.0486	0.0603	0.0700	0.0775	0.0877	0.0968	**Green**
Skew-*t*-GJR-M	0.0110	0.0204	0.0301	0.0392	0.0494	0.0594	0.0701	0.0775	0.0882	0.0972	**Green**
					NASDAQ						
normal-GARCH	0.0112	0.0230	0.0354	0.0453	0.0553	0.0639	0.0742	0.0840	0.0928	0.1055	**Green**
t-GARCH	0.0116	0.0234	0.0338	0.0448	0.0543	0.0643	0.0744	0.0842	0.0935	0.1047	**Green**
Skew-*t*-GARCH	0.0116	0.0225	0.0332	0.0440	0.0536	0.0637	0.0727	0.0824	0.0923	0.1041	**Green**
normal-GARCH-M	0.0107	0.0193	0.0310	0.0411	0.0496	0.0570	0.0656	0.0743	0.0836	0.0935	**Green**
t-GARCH-M	0.0103	0.0203	0.0315	0.0408	0.0501	0.0577	0.0668	0.0759	0.0858	0.0963	**Green**
Skew-*t*-GARCH-M	0.0101	0.0202	0.0314	0.0408	0.0503	0.0584	0.0671	0.0763	0.0848	0.0962	**Green**
normal-GJR	0.0121	0.0260	0.0364	0.0460	0.0570	0.0661	0.0775	0.0874	0.0968	0.1069	**Yellow**
t-GJR	0.0116	0.0250	0.0361	0.0460	0.0567	0.0656	0.0774	0.0873	0.0969	0.1075	**Green**
Skew-*t*-GJR	0.0115	0.0242	0.0353	0.0441	0.0562	0.0645	0.0760	0.0853	0.0950	0.1054	**Green**
normal-GJR-M	0.0113	0.0230	0.0340	0.0431	0.0531	0.0611	0.0706	0.0806	0.0883	0.0993	**Green**
t-GJR-M	0.0110	0.0228	0.0346	0.0428	0.0518	0.0599	0.0701	0.0808	0.0908	0.0995	**Green**
Skew-*t*-GJR-M	0.0116	0.0228	0.0336	0.0421	0.0527	0.0610	0.0705	0.0800	0.0893	0.0983	**Green**
					NIKKEI						
normal-GARCH	0.0106	0.0210	0.0311	0.0411	0.0524	0.0618	0.0721	0.0825	0.0929	0.1029	**Green**
t-GARCH	0.0105	0.0208	0.0321	0.0423	0.0515	0.0620	0.0731	0.0835	0.0919	0.1028	**Green**
Skew-*t*-GARCH	0.0110	0.0203	0.0318	0.0423	0.0519	0.0611	0.0735	0.0836	0.0933	0.1025	**Green**
normal-GARCH-M	0.0098	0.0195	0.0273	0.0371	0.0461	0.0560	0.0648	0.0746	0.0827	0.0919	**Green**
t-GARCH-M	0.0104	0.0186	0.0283	0.0370	0.0461	0.0565	0.0659	0.0764	0.0842	0.0925	**Green**
Skew-*t*-GARCH-M	0.0101	0.0184	0.0283	0.0380	0.0470	0.0572	0.0666	0.0770	0.0839	0.0924	**Green**
normal-GJR	0.0113	0.0219	0.0326	0.0421	0.0528	0.0626	0.0736	0.0822	0.0926	0.1041	**Green**
t-GJR	0.0116	0.0211	0.0330	0.0422	0.0529	0.0627	0.0744	0.0824	0.0932	0.1047	**Green**
Skew-*t*-GJR	0.0118	0.0210	0.0324	0.0430	0.0528	0.0628	0.0742	0.0823	0.0929	0.1046	**Yellow**
normal-GJR-M	0.0118	0.0224	0.0306	0.0406	0.0493	0.0599	0.0694	0.0767	0.0867	0.0972	**Yellow**
t-GJR-M	0.0119	0.0206	0.0302	0.0408	0.0492	0.0598	0.0699	0.0788	0.0876	0.0983	**Yellow**
Skew-*t*-GJR-M	0.0123	0.0210	0.0306	0.0414	0.0494	0.0608	0.0695	0.0790	0.0877	0.0991	**Yellow**
					S&P500						
normal-GARCH	0.0124	0.0223	0.0334	0.0435	0.0524	0.0617	0.0716	0.0800	0.0908	0.0996	**Yellow**
t-GARCH	0.0123	0.0227	0.0339	0.0429	0.0527	0.0618	0.0720	0.0806	0.0912	0.1005	**Yellow**
Skew-*t*-GARCH	0.0127	0.0222	0.0333	0.0433	0.0529	0.0620	0.0709	0.0818	0.0910	0.1013	**Yellow**
normal-GARCH-M	0.0110	0.0199	0.0288	0.0383	0.0465	0.0532	0.0611	0.0698	0.0791	0.0868	**Green**
t-GARCH-M	0.0111	0.0207	0.0306	0.0392	0.0476	0.0546	0.0644	0.0732	0.0829	0.0905	**Green**
Skew-*t*-GARCH-M	0.0110	0.0209	0.0307	0.0397	0.0472	0.0544	0.0642	0.0735	0.0831	0.0908	**Green**
normal-GJR	0.0141	0.0234	0.0334	0.0444	0.0538	0.0623	0.0722	0.0821	0.0914	0.1006	**Red**
t-GJR	0.0127	0.0232	0.0326	0.0440	0.0526	0.0625	0.0718	0.0811	0.0918	0.1008	**Yellow**
Skew-*t*-GJR	0.0125	0.0237	0.0330	0.0435	0.0528	0.0628	0.0713	0.0812	0.0919	0.1022	**Yellow**
normal-GJR-M	0.0116	0.0209	0.0292	0.0371	0.0462	0.0542	0.0609	0.0696	0.0785	0.0855	**Green**
t-GJR-M	0.0116	0.0214	0.0303	0.0395	0.0490	0.0561	0.0641	0.0728	0.0812	0.0910	**Green**
Skew-*t*-GJR-M	0.0119	0.0212	0.0305	0.0391	0.0488	0.0557	0.0635	0.0729	0.0824	0.0916	**Yellow**

* Classification according to the three-zone appraoch suggested by the Basle Committee (1996): a VaR model is accurate (green zone) if the number of violations of 1% VaR remains below the binomial (0.01) 95% quantile. A model is arguable (yellow zone) up to 99.99% quantile. When more violations occur the model is judged as inappropriate (red zone). For our sample size, if at most 109 (1.17%) violations occur the model is acceptable. Between 110 and 130 (1.39%) is disputable.

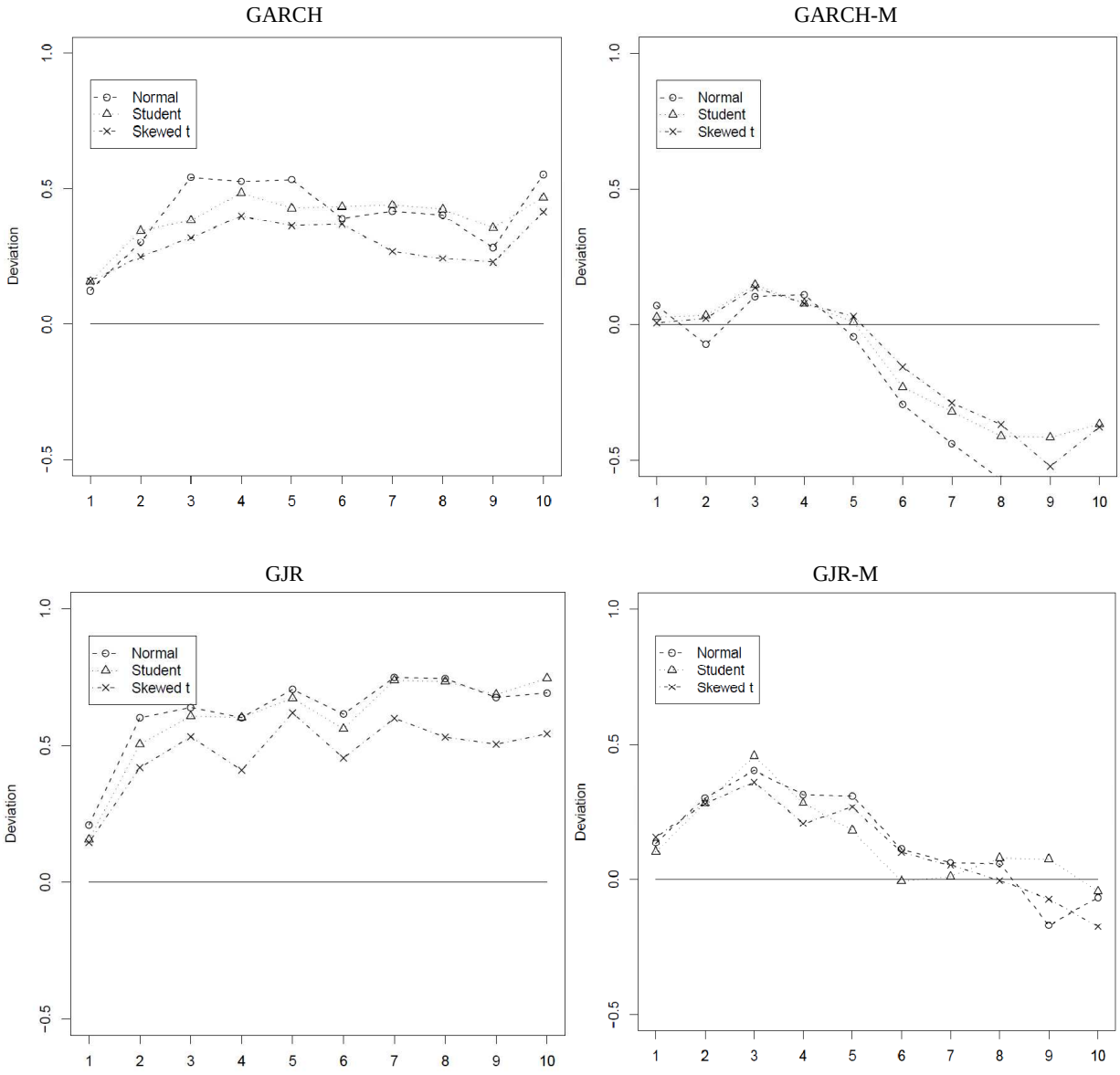

Figure 1 Deviation probability plot for the FHS filtered by GARCH-type models. The horizontal axis is the VaR level. In the vertical axis, for each VaR level, the excess of percentage violations over the VaR level is represented.

Table 9 Overall measures of deviation

	normal		*t* -Student		Skewed-*t*	
	MAE (%)	MSE(%)	MAE (%)	MSE(%)	MAE (%)	MSE(%)
GARCH	0.2061	0.0006	0.2043	0.0006	0.1953	0.0005
GARCH-M	0.3911	0.0026	0.2853	0.0013	0.2788	0.0013
GJR	0.2979	0.0013	0.2986	0.0012	0.2734	0.0010
GJR-M	0.2724	0.0016	0.2126	0.0009	0.2066	0.0008

According to the reported results for the MAE and MSE, the traditional normal-GARCH model is the most appropriate model to filter stock returns in the context of FHS. Also, we can conclude that the use of alternative innovation assumptions (*t*-Student and Skewed-*t*) has no impact on the VaR forecasting performance of FHS coupled with a GARCH model.

We should now attempt on the information in the sequence of violations, provided by the P-values of the LR and DQ test statistics described in section 3.2. As an example, the detailed results of the VaR forecast performance for the NASDAQ sample, are presented in table 10[8]. Table 11 summarizes the number of rejections of the null hypothesis by applying the three LR tests and the DQ test to all the tested models across the six empirical samples.

According to the results, when a traditional normal-GARCH model is used to filter the returns in order to estimate VaR by FHS, a poor performance in terms of independence is reported. When instead a GJR process is used, there is a substantial improvement with respect to the independence of the VaR violations. In fact, the results of the DQ tests clearly indicate that the use of a GJR, improves the VaR forecasting performance of FHS in terms of independence. It should be noticed that the use of alternative distributional assumptions for the innovations, by itself, does not have an important impact in the results.

[8] In the interest of brevity, the detailed results for the VaR forecast performance with respect to the other five empirical samples are not included in the paper. These results are available upon request.

Table 10 VaR forecast performance: NASDAQ[a]

Target downfall probability	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
				P_{uc}						
normal-GARCH	**0.2381**	**0.0422**	0.0028	**0.0108**	**0.0199**	**0.1162**	**0.1172**	**0.1551**	**0.3471**	**0.0770**
t -GARCH	**0.1393**	**0.0207**	**0.0337**	**0.0190**	**0.0619**	**0.0816**	**0.0997**	**0.1345**	**0.2338**	**0.1345**
Skew-*t* -GARCH	**0.1393**	**0.0935**	**0.0763**	**0.0526**	**0.1121**	**0.1374**	**0.3137**	**0.3918**	**0.4467**	**0.1847**
normal-GARCH-M	**0.4994**	**0.6100**	**0.5581**	**0.5889**	**0.8436**	**0.2259**	**0.0931**	**0.0393**	**0.0287**	**0.0357**
t -GARCH-M	**0.7902**	**0.8170**	**0.4094**	**0.7016**	**0.9678**	**0.3442**	**0.2203**	**0.1395**	**0.1570**	**0.2343**
Skew-*t* -GARCH-M	**0.9528**	**0.8746**	**0.4441**	**0.7016**	**0.8925**	**0.5236**	**0.2703**	**0.1859**	**0.0747**	**0.2209**
normal-GJR	**0.0488**	0.0001	0.0005	0.0037	0.0022	**0.0138**	0.0052	0.0089	**0.0240**	**0.0272**
t -GJR	**0.1393**	0.0008	0.0009	0.0037	0.0035	**0.0244**	0.0059	0.0099	**0.0219**	**0.0174**
Skew-*t* -GJR	**0.1679**	0.0051	0.0033	**0.0467**	0.0070	**0.0677**	**0.0250**	**0.0613**	**0.0911**	**0.0828**
normal-GJR-M	**0.2007**	**0.0422**	**0.0251**	**0.1267**	**0.1750**	**0.6508**	**0.8084**	**0.8323**	**0.5651**	**0.8278**
t -GJR-M	**0.3276**	**0.0588**	**0.0114**	**0.1700**	**0.4264**	**0.9798**	**0.9680**	**0.7735**	**0.7969**	**0.8820**
Skew-*t* -GJR-M	**0.1393**	**0.0588**	**0.0448**	**0.3127**	**0.2420**	**0.6823**	**0.8398**	**0.9866**	**0.8036**	**0.5730**
				P_{ind}						
normal-GARCH	**0.4769**	**0.0902**	0.0050	**0.1688**	**0.4589**	**0.3678**	**0.5810**	**0.3657**	**0.3097**	**0.5235**
t -GARCH	**0.5191**	**0.1095**	0.0035	**0.0581**	**0.1905**	**0.5893**	**0.1999**	**0.2713**	**0.2775**	**0.2118**
Skew-*t* -GARCH	**0.5191**	**0.1539**	**0.0109**	**0.0626**	**0.1496**	**0.2095**	**0.1019**	**0.1105**	**0.1787**	**0.1444**
normal-GARCH-M	**0.4095**	**0.2022**	**0.0131**	**0.0418**	**0.4647**	**0.4469**	**0.3057**	**0.3025**	**0.2684**	**0.3279**
t -GARCH-M	**0.3587**	**0.0216**	**0.0692**	**0.0609**	**0.2255**	**0.2513**	**0.2745**	**0.1287**	**0.3069**	**0.0847**
Skew-*t* -GARCH-M	**0.3345**	**0.0202**	**0.0160**	**0.1007**	**0.1724**	**0.3183**	**0.0746**	**0.1175**	**0.1750**	**0.0215**
normal-GJR	**0.2172**	**0.7214**	**0.1130**	**0.5506**	**0.4532**	**0.6787**	**0.2061**	**0.1145**	**0.0700**	**0.1361**
t -GJR	**0.1752**	**0.3737**	**0.1641**	**0.5506**	**0.5966**	**0.6725**	**0.2705**	**0.6465**	**0.3063**	**0.3252**
Skew-*t* -GJR	**0.1674**	**0.2871**	**0.1215**	**0.3368**	**0.6550**	**0.7147**	**0.5533**	**0.6581**	**0.5432**	**0.3220**
normal-GJR-M	**0.1599**	**0.1897**	**0.4746**	**0.6187**	**0.7358**	**0.7221**	**0.4047**	**0.1126**	**0.1056**	**0.3390**
t -GJR-M	**0.1388**	**0.0817**	**0.2448**	**0.5814**	**0.4915**	**0.6614**	**0.3023**	**0.6183**	**0.5061**	**0.4825**
Skew-*t* -GJR-M	**0.1752**	**0.1748**	**0.1009**	**0.1697**	**0.2810**	**0.5941**	**0.6964**	**0.6818**	**0.5334**	**0.4727**
				P_{cc}						
normal-GARCH	**0.3872**	**0.0303**	0.0002	**0.0150**	**0.0506**	**0.1941**	**0.2517**	**0.2418**	**0.3837**	**0.1708**
t -GARCH	**0.2724**	**0.0192**	0.0015	**0.0106**	**0.0743**	**0.1897**	**0.1134**	**0.1781**	**0.2729**	**0.1496**
Skew-*t* -GARCH	**0.2724**	**0.0887**	0.0082	**0.0270**	**0.1002**	**0.1509**	**0.1580**	**0.1939**	**0.3031**	**0.1430**
normal-GARCH-M	**0.5665**	**0.3893**	**0.0389**	**0.1088**	**0.7507**	**0.3597**	**0.1445**	**0.0703**	**0.0495**	**0.0683**
t -GARCH-M	**0.6334**	**0.0696**	**0.1366**	**0.1604**	**0.4794**	**0.3311**	**0.2597**	**0.1058**	**0.2180**	**0.1116**
Skew-*t* -GARCH-M	**0.6266**	**0.0667**	**0.0411**	**0.2415**	**0.3906**	**0.4959**	**0.1111**	**0.1224**	**0.0814**	**0.0336**
normal-GJR	**0.0671**	0.0004	0.0006	**0.0123**	0.0070	**0.0442**	0.0091	0.0094	**0.0152**	**0.0287**
t -GJR	**0.1337**	0.0024	0.0015	**0.0123**	**0.0121**	**0.0725**	**0.0123**	**0.0323**	**0.0428**	**0.0365**
Skew-*t* -GJR	**0.1490**	**0.0112**	0.0041	**0.0872**	**0.0240**	**0.1763**	**0.0681**	**0.1574**	**0.1994**	**0.1360**
normal-GJR-M	**0.1643**	**0.0538**	**0.0630**	**0.2753**	**0.3765**	**0.8473**	**0.6862**	**0.2777**	**0.2288**	**0.6183**
t -GJR-M	**0.2071**	**0.0368**	**0.0207**	**0.3350**	**0.5753**	**0.9083**	**0.5870**	**0.8474**	**0.7755**	**0.7729**
Skew-*t* -GJR-M	**0.1337**	**0.0667**	**0.0348**	**0.2340**	**0.2821**	**0.7979**	**0.9080**	**0.9193**	**0.7986**	**0.6593**
				DQ_{Hit}						
normal-GARCH	0.0073	**0.0116**	0.0000	0.0012	0.0078	**0.0507**	**0.2812**	**0.0428**	0.0065	**0.0146**
t -GARCH	**0.0461**	0.0001	0.0000	0.0013	**0.0102**	**0.0648**	**0.0194**	**0.0460**	0.0008	0.0004
Skew-*t* -GARCH	**0.0461**	0.0054	0.0002	0.0032	0.0050	0.0074	**0.0115**	**0.0208**	0.0018	0.0001
normal-GARCH-M	0.0045	**0.0859**	0.0030	**0.0232**	**0.0201**	**0.0151**	**0.0259**	0.0050	**0.0108**	0.0016
t -GARCH-M	**0.2665**	0.0029	0.0019	**0.0335**	**0.0109**	**0.0241**	**0.0252**	**0.0203**	**0.0607**	0.0001
Skew-*t* -GARCH-M	**0.0720**	**0.0139**	0.0019	**0.0480**	**0.0233**	**0.0140**	0.0038	**0.0108**	0.0009	0.0000
normal-GJR	**0.1909**	0.0027	0.0046	**0.0811**	**0.0136**	**0.1598**	**0.0755**	**0.0671**	**0.0894**	**0.0833**
t -GJR	**0.1231**	0.0041	0.0072	**0.0617**	**0.0138**	**0.1643**	**0.0868**	**0.1795**	**0.1904**	**0.0786**
Skew-*t* -GJR	**0.2882**	**0.0211**	0.0055	**0.2327**	0.0080	**0.2585**	**0.2804**	**0.4629**	**0.3369**	**0.1220**
normal-GJR-M	**0.1117**	**0.1510**	**0.2835**	**0.7060**	**0.1668**	**0.6045**	**0.9697**	**0.7612**	**0.5729**	**0.6806**
t -GJR-M	**0.1276**	0.0031	**0.0514**	**0.4763**	**0.5578**	**0.4554**	**0.7375**	**0.9669**	**0.8467**	**0.2981**
Skew-*t* -GJR-M	**0.2756**	**0.0276**	**0.0105**	**0.1362**	**0.1616**	**0.5066**	**0.7221**	**0.9154**	**0.3055**	**0.1050**
				DQ_{VaR}						
normal-GARCH	0.0004	0.0027	0.0000	0.0008	0.0097	**0.0579**	**0.1801**	**0.0263**	0.0039	**0.0119**
t -GARCH	0.0026	0.0000	0.0000	0.0004	0.0088	**0.0538**	0.0092	**0.0261**	0.0004	0.0004
Skew-*t* -GARCH	0.0058	0.0012	0.0000	0.0013	0.0056	**0.0103**	**0.0101**	**0.0132**	0.0016	0.0001
normal-GARCH-M	0.0002	**0.0202**	0.0005	0.0063	**0.0268**	**0.0220**	**0.0346**	0.0063	**0.0183**	0.0027
t -GARCH-M	**0.0435**	0.0007	0.0001	**0.0135**	**0.0110**	**0.0389**	**0.0257**	**0.0187**	**0.0736**	0.0002
Skew-*t* -GARCH-M	**0.0101**	0.0035	0.0008	**0.0211**	**0.0305**	**0.0232**	0.0048	**0.0109**	0.0014	0.0000
normal-GJR	**0.0102**	0.0012	0.0016	**0.0222**	**0.0141**	**0.1538**	**0.1052**	**0.0842**	**0.1169**	**0.0783**
t -GJR	0.0072	0.0013	0.0009	**0.0147**	0.0098	**0.1667**	**0.0777**	**0.1340**	**0.1467**	**0.0478**
Skew-*t* -GJR	**0.0453**	0.0065	0.0014	**0.0978**	0.0075	**0.2631**	**0.2917**	**0.4763**	**0.4030**	**0.1605**
normal-GJR-M	0.0088	**0.0316**	**0.0721**	**0.2818**	**0.1580**	**0.5155**	**0.9595**	**0.8220**	**0.6909**	**0.6984**
t -GJR-M	**0.0111**	0.0002	0.0052	**0.1324**	**0.3037**	**0.3606**	**0.6996**	**0.9417**	**0.6898**	**0.2184**
Skew-*t* -GJR-M	**0.0354**	0.0092	0.0023	**0.0238**	**0.0599**	**0.3226**	**0.6863**	**0.9003**	**0.3302**	**0.1210**

[a] Entries in the last 10 columns are the P-values of the respective tests. Bold type entries are not significant at the 1% level.

Table 11 Number of rejections of the null hypothesis (number of P-values below the desired level)

	sig. Level 0.01									
	P_{uc}	P_{ind}	P_{cc}	DQ_{Hit}	DQ_{VaR}	P_{uc}	P_{ind}	P_{cc}	DQ_{Hit}	DQ_{VaR}
	normal-GARCH					normal-GJR				
DAX	0	0	0	5	5	0	0	0	0	0
DJI	0	1	2	2	4	0	1	2	2	3
FTSE	0	2	0	0	0	0	0	0	0	0
NASDAQ	1	1	1	5	6	6	0	5	2	2
Nikkei	0	0	0	3	4	0	0	0	0	0
S&P 500	0	2	1	4	8	1	1	2	3	4
	t-GARCH					*t*-GJR				
DAX	0	0	0	9	9	0	0	0	0	1
DJI	0	2	2	3	4	0	2	2	2	3
FTSE	0	4	0	0	2	0	0	0	0	0
NASDAQ	0	1	1	5	8	6	0	2	2	4
Nikkei	0	0	0	10	10	0	0	0	0	0
S&P 500	0	5	3	8	9	0	1	2	3	4
	Skew-*t*-GARCH					Skew-*t*-GJR				
DAX	0	0	0	9	9	0	0	0	0	0
DJI	0	3	2	3	5	1	2	2	2	3
FTSE	0	4	0	0	3	0	0	0	0	0
NASDAQ	0	0	1	7	7	3	0	1	2	3
Nikkei	0	1	0	9	10	0	0	0	0	1
S&P 500	0	5	4	8	9	0	0	1	1	4
	normal-GARCH-M					normal-GJR-M				
DAX	1	0	0	5	5	0	0	0	0	1
DJI	4	2	6	7	7	1	3	0	2	10
FTSE	0	3	1	0	1	0	0	0	0	0
NASDAQ	0	0	0	4	5	0	0	0	0	1
Nikkei	1	0	1	3	5	0	0	0	0	2
S&P 500	5	3	6	8	9	4	0	4	4	6
	t-GARCH-M					*t*-GJR-M				
DAX	0	0	0	7	7	0	0	0	0	0
DJI	0	5	5	5	9	0	2	0	0	6
FTSE	0	4	1	1	2	0	0	0	0	0
NASDAQ	0	0	0	3	3	0	0	0	1	2
Nikkei	0	0	1	9	9	0	0	0	0	1
S&P 500	1	4	5	8	9	3	2	2	2	6
	Skew-*t*-GARCH-M					Skew-*t*-GJR-M				
DAX	0	0	0	7	7	0	0	0	0	0
DJI	0	4	3	5	7	0	2	1	1	3
FTSE	0	5	2	3	4	0	0	0	1	2
NASDAQ	0	0	0	4	5	0	0	0	0	2
Nikkei	0	1	1	9	8	0	0	0	0	2
S&P 500	1	2	5	9	9	2	2	1	1	3

5. Concluding remarks

The predictive performance of FHS combined with a traditional normal-GARCH model has been backtested using simulated data. Three realistic different DGP were considered to generate several series of simulated returns. Based on the simulated series, a backtesting was implemented. Our backtesting results are very promising as they indicate the validity of FHS to forecast VaR, with respect to the three alternative DGPs. For the great majority of the samples, the null hypothesis that the risk model is correct on average is not rejected. Additionally, two competing models, differing in the innovations assumption, were tested: a *t*-GARCH and a Skew-*t*-GARCH. The results indicate that the use of alternative innovations assumptions does not generally impacts the FHS results in terms of VaR forecasting performance. In summary, the simulation results strongly indicates that FHS is an accurate method (in terms of coverage and independence) to forecast VaR in the presence of non-normal returns. Moreover, our results demonstrate that FHS can be applied to forecast VaR for data which exhibits a high incidence of zeros, time-varying skewness, asymmetric effects to return shocks on volatility as well as other stylized facts.

Though the use of simulated data enable us to assess the VaR forecasting ability of FHS under controllable circumstances, the validity of any method is best measured using empirical time series. Six well known active stock indices daily time series were used to produce the empirical results. The VaR forecasting ability of FHS method, using four competitive GARCH-type models to filter the stock returns, combined with three alternative innovation assumptions, was tested. Though all the models demonstrate a good performance, according to our empirical coverage results the traditional normal-GARCH model is the most appropriate model to filter stock returns in the context of FHS. Also, we have concluded that the results are not sensitive to the use of alternative innovation assumptions (*t*-Student and Skewed-*t*). Nevertheless, when the VaR forecast performance of FHS is assessed in terms of independence, some problems are reported for the traditional normal-GARCH model. With respect to the empirical data series, the choice of a GJR process results in a substantial improvement

with respect to the independence of the VaR violations. In fact, the results of the DQ tests clearly indicate that the use of a GJR, improves the VaR forecasting performance of FHS. It should be noticed that, again, the use of alternative distributional assumptions for the innovations does not have an important impact in the results. This is a very important result as it gives support for the argument that distributionally nonparametric models, like FHS, do not depend on the distribution assumed in the filtering stage.

References

Alexander, C. and Lazar, E. (2006) *Normal Mixture GARCH (1,1): Applications to Foreign Exchange Markets.* Journal of Applied Econometrics, 21(3), 307-336.

Barone-Adesi, G., Giannopoulos, K. and Vosper, L. (1999) *VaR Without Correlations for Portfolio of Derivative Securities.* Journal of Futures Markets, 19, 583-602.

Barone-Adesi, G., Giannopoulos, K. and Vosper, L. (2002). *Backtesting Derivative Portfolios with FHS*. European Financial Management, 8, 31-58.

Basle Committee on Banking Supervision (1996). *Supervisory Framework for the use of Backtesting in Conjunction with Internal Models Approach to Market Risk Capital Requirements*. Available at http://bis.org.

Bucay, N. and Rosen, D. (1999). *Credit Risk of an International Bond Portfolio: A Case Study*. Algo Research Quarterly, 1(2), 9-29.

Campbell, J. Y. and Hentschel, L. (1992). *No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns*, Journal of Financial Economics, 31, 281-318.

Christofferson, P. F. (2003) *Elements of Financial Risk Management*. Academic Press. London.

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. (2007) *Object-Oriented Matrix Programming Using Ox*. 3rd ed. London: Timberlake Consultants Press and Oxford: www.doornik.com.

Engle, R. F., Lilien, D. and Robins, R. (1987). *Estimation of Time Vaying Risk Premiums in the Term Structure*, Econometrica, 55, 391-408

Engle, R.F. and Manganelli S. (2004). *CAViaR: Conditional autoregressive Value at Risk by regression quantile*, Journal of Business and Economic Statistics, 22, 367-381.

Glosten, L. R., Jagannathan, R. and Runkle, D. E. (1993). *On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks*, Journal of Finance, American Finance Association, December, 48(5), 1779-1801.

Hartz, C., Mittnik, S. and Paolella, M. (2006). *Accurate Value-at-Risk Forecasting Based on the Normal-GARCH Model*. Computational Statistics & Data Analysis, 51, 2295-2312.

Haas, M., Mittnik, S. and Paolella, M. S. (2004). *A New Approach to Markov-Switching GARCH Models*. Journal of Financial Econometrics, 2(4), 493-530.

Mausser, H. and Rosen, D. (1999). *Beyond VaR: From Measuring Risk to Managing Risk*. ALGO Research Quarterly, 1(2), 5-20.

Mittnik, S. and Paolella, M.S. (2000) *Conditional density and value-at-risk prediction of Asian currency exchange rates*. Journal of Forecast, 19, 313-333.

Paolella, M. S. and Taschini L. (2008) *An Econometric Analysis of Emission Allowances Prices*, Journal of Banking and Finance, 32 (10), 2022-2032.

Pritsker, M. (2001). *The Hidden Dangerous of Historical Simulation*. Working Paper, January. Federal Reserve Board.

Kuester K., Mittinik S. and Paolella M. S. (2005) *Value-at-Risk Prediction: A Comparison of Alternative Strategies*, Journal of Financial Econometrics, 4 (1), 53-89.