HIAS-E-16

Employing Bayesian Forecasting of Value-at-Risk to Determine an Appropriate Model for Risk Management

Cathy W.S. Chen, Monica M.C. Weng
Graduate Institute of Statistics & Actuarial Science, Feng Chia University, Taiwan
Toshiaki Watanabe
Institute of Economic Research, Hitotsubashi University

December 08, 2015

Hitotsubashi Institute for Advanced Study, Hitotsubashi University
2-1, Naka, Kunitachi, Tokyo 186-8601, Japan
tel:+81 42 580 8604   http://hias.ad.hit-u.ac.jp/

Employing Bayesian Forecasting of Value-at-Risk to Determine an Appropriate Model for Risk Management

Cathy W.S. Chen,*[a] Monica M.C. Weng,[a] and Toshiaki Watanabe[b]

[a]Graduate Institute of Statistics & Actuarial Science, Feng Chia University, Taiwan

[b]Institute of Economic Research, Hitotsubashi University, Japan

Abstract

To allow for a higher degree of flexibility in model parameters, we propose a general and time-varying nonlinear smooth transition (ST) heteroskedastic model with a second-order logistic function of varying speed in the mean and variance. This paper evaluates the performance of Value-at-Risk (VaR) measures in a class of risk models, specially focusing on three distinct ST functions with GARCH structures: first- and second-order logistic functions, and the exponential function. The likelihood function is non-differentiable in terms of the threshold values and delay parameter. We employ Bayesian Markov chain Monte Carlo sampling methods to update the estimates and quantile forecasts. The proposed methods are illustrated using simulated data and an empirical study. We estimate VaR forecasts for the proposed models alongside some competing asymmetric models with skew and fat-tailed error probability distributions, including realized volatility models. To evaluate the accuracy of VaR estimates, we implement two loss functions and three backtests. The results show that the ST model with a second-order logistic function and skew Student's t error is a worthy choice at the 1% level, when compared to a range of existing alternatives.

Keywords: Second-order logistic transition function; Backtesting; Markov chain Monte Carlo methods; Value-at-Risk; Volatility forecasting; Realized volatility models.

*Corresponding author: Cathy W.S. Chen, Email:chenws@mail.fcu.edu.tw

1 Introduction

Financial risk management is widely used in financial institutions in order to control risk exposures, such as credit risk, operation risk, and volatility. The Basel II Accord, initially published in June 2004, is intended to create international standards for banking regulators to better control risk exposures. In theory, Basel II set up risk and capital management requirements designed to ensure that Authorized Deposit-taking Institutions (ADIs) have enough financial ability to maintain solvency. One of the widely used risk measures is Value-at-risk (VaR), which is designed to forecast the worst expected loss over a given time interval under normal market conditions, at a given confidence level α (Jorion 1997). The financial firm JP Morgan proposed VaR for reporting firm wide risk in its RiskMetrics model in 1993, known as the famous "4:15 pm report." By 1996, amendments to the Basel II Accord permitted banks to use an "appropriate model" to calculate their VaR thresholds.

There are many VaR estimation methods in the literature that can be classified into three categories. They include non-parametric methods, for example, historical simulation (HS) (using past or in-sample quantiles); semi-parametric methods, for example, the extreme value theory, the dynamic quantile regression CAViaR model (Engle & Manganelli, 2004), and the threshold CAViaR model (Gerlach Chen, & Chan, 2011, Chen et al., 2012); and parametric statistical approaches that fully specify model dynamics and distribution assumptions, for example, the autoregressive conditional heteroskedasticity (ARCH) model proposed by Engle (1982) and its generalized version by Bollerslev (1986), popularly known as the GARCH model. It is well known that the GARCH model cannot capture the asymmetric response of volatility, the phenomenon discovered by Rabemananjara and Zakoïan (1993) and Zakoïan (1994), among others.

Chan and Tong (1986) introduce a smooth transition (ST) autoregressive model to allow for flexibility in model parameters through a smooth transition, which gained popularity via Granger and Teräsvirta (1993) and Teräsvirta (1994). Their first-order logistic function gives a continuous value between zero and one. Jansen and Teräsvirta (1996) appear to be the first to discuss the second-order logistic function in ST models. Van Dijk, Teräsvirta, and Franses (2002) investigate the second-order logistic function with a slight difference in format from that of Jansen and Teräsvirta (1996), but these two papers only focus on a transition in describing the mean equation. Most financial time series exhibit an asymmetric behavior in the mean and in the volatility as well. The 2-regime ST models can be extended to allow for more than two regimes and a time-varying smooth transition conditional variance.

In this paper we propose a more general and time-varying ST-GARCH model to allow an ST function with varying speed in the mean and variance. The second-order logistic function has been described in the literature, but not often used in practice due to the difficulty in parameter estimation. The problem of estimating the second-order ST-GARCH models has become a challenge. Specifically, the likelihood function is non-differentiable in terms of the threshold values and delay parameter. Our paper examines whether such double asymmetry might be better modeled by an ST function in both the mean and volatility equations for VaR forecasting and volatility estimation. Based on Markov chain Monte Carlo (MCMC) methods, we employ a Bayesian approach and allow a simultaneous inference for all unknown parameters, while the parameter constraints simply and properly form part of the prior distribution, and the problems with estimating threshold limits and delay lags disappear. To our knowledge, this study is the first in the literature to make Bayesian inferences and quantile forecasting for the ST-GARCH model with a second-order logistic function.

We employ three distinct ST functions with autoregressive conditional heteroskedastic models for VaR forecasting purposes: the first- and second-order logistic functions, and the exponential functions. Gerlach and Chen (2008) incorporate the first-order ST functions into GARCH models to allow for smooth nonlinearity in the mean and asymmetry of the volatility. Chen et al. (2010) employ an exponential function to capture size asymmetry in the mean and volatility. Compared to existing models, our proposed model conveys that observations in the extremes can have a dissimilar effect and an ST function with varying speed in the mean and variance. Moreover, the ST-GARCH model, with the second-order logistic function, can be viewed as three regimes interpreted as follows: the first regime is related to extremely low negative shocks ("bad news"), the middle regime represents low absolute returns ("tranquil periods"), and finally, the third regime is related to high positive shocks ("good news").

As discussed in Andersen, Bollerslev, Diebold and Labys (2001), the theory of quadratic variation indicates that, under suitable conditions, realized volatility is an unbiased and highly efficient estimator of return volatility. We also deal with the inclusion of realized measures of volatility in a GARCH modelling setup. The realized GARCH (RV) model of Hansen, Huang and Shek (2012) provides an excellent framework for the joint modelling of returns and realized measures of volatility.

This paper focuses on parametric models and Monte Carlo simulation to forecast VaR. We consider popular variants and extensions of the GARCH model family as follows: RiskMetrics; GARCH; asymmetric GJR-GARCH (Glosten, Jaganathan, Runkle; Glosten et al., 1993); ST-GARCH with three distinct ST functions; and threshold nonlinear GARCH (TGARCH; Chen and So 2006). Each model includes a specification for the volatility dynamics, and most consider three specifications for the conditional asset return distribution: Gaussian, Student's t, and the skew Student's t of Hansen (1994). This paper extensively examines the VaR forecast performance over 12 risk models and two HS methods during two out-of-sample periods: the two-year post-global financial crisis period and the three-year global financial crisis period. To shed light on the advantage of Bayesian updating forecasting, this study examines a sample of ten European stock markets, seven Asian stock markets, one North American market, and one South American market, for a total of 19 stock markets over the post-global financial crisis period. We focus on Japan and U.S. stock markets for the three-year global financial crisis period, including the RV model.

The use of our proposed Bayesian forecasting of nonlinear ST models to deal with some complex derivatives and to calculate their corresponding VaR formulae is of practical importance and theoretical interests. Bayesian MCMC methods have many advantages in estimation, inference, and forecasting, including: (i) accounting for parameter uncertainty in both probabilistic and point forecasting; (ii) allowing simultaneous inference for all unknown parameters; (iii) efficient and flexible handling of complex models and nonstandard parameters; and (iv) parameter constraints simply and properly form part of the prior distribution. As such, MCMC methods are generally used to forecast VaR thresholds for each risk model in this paper. We follow the procedure in Chen and So (2006) and design an adaptive MCMC sampling scheme for estimation and quantile forecasting.

When forecasting VaR thresholds, our aim is to find the optimal combination of volatility dynamics and error distribution in terms of the observed violation rates and two loss functions in Lopez (1999 a, b) for both out-of-sample periods. We further consider three backtesting methods for evaluating and testing the accuracy of VaR models. We also investigate the accuracy of volatility forecasts for all models under three volatility proxies with three loss functions.

This paper is organized as follows. Section 2 illustrates the ST models with different ST functions. Section 3 demonstrates the Bayesian setup and details of parameter inferences. Section 4 describes the process of VaR forecasting. Section 5 presents a simulation study of a double ST-GARCH model with the second-order logistic function showing the estimation performance. We further extend this class of ST-GARCH models to incorporate a different effect (smooth transition function) for the mean and variance. Section 6 presents empirical results, focusing on the forecasts of VaR and volatility and furthermore showing the forecast accuracy for all models under three volatility proxies. Section 7 provides concluding remarks.

2 The smooth transition heteroskedastic model

We consider a general double smooth transition GARCH model to capture mean and volatility asymmetry in financial markets. We present the ST-GARCH model as below:

$$
\begin{aligned}
y_t &= \mu_t^{(1)} + F(z_{t-d};\gamma,c)\mu_t^{(2)} + a_t \\
a_t &= \sqrt{h_t}\epsilon_t, \qquad \epsilon_t \overset{\text{i.i.d.}}{\sim} D(0,1), \\
h_t &= h_t^{(1)} + F(z_{t-d};\gamma,c)h_t^{(2)} \\
\mu_t^{(l)} &= \phi_0^{(l)} + \sum_{i=1}^{p}\phi_i^{(l)} y_{t-i} \\
h_t^{(l)} &= \alpha_0^{(l)} + \sum_{i=1}^{g}\alpha_i^{(l)} a_{t-i}^2 + \sum_{i=1}^{q}\beta_i^{(l)} h_{t-i}, \quad l = 1,2,
\end{aligned}
\tag{1}
$$

where z_t is the threshold variable; d is the delay lag; and $D(0,1)$ is an error distribution with mean 0 and variance 1. The parameter γ determines the smoothness of the change in the value of $F(z_{t-d};\gamma,c)$ function and the smoothness of the transition from one regime to the other. We consider three types of ST functions in this work. Different choices for the transition function lead to different types of regime-switching behaviour. A popular choice for $F(z_{t-d};\gamma,c)$ is the first-order logistic function:

$$F(z_{t-d};\gamma,c) = \frac{1}{1+\exp\left\{\frac{-\gamma(z_{t-d}-c)}{s_z}\right\}}, \tag{2}$$

where s_z is the sample standard deviation of z_t. This type of regime-switching can be convenient for modelling, for example, asymmetry in stock markets to distinguish bad news and good news. The first-order logistic ST is an odd function and is used to capture sign asymmetry; in other words, the asymmetric responses to positive and negative values of $z_{t-d} - c$. Teräsvirta and Anderson (1992) and Teräsvirta (1994) apply the STAR model with a first-order logistic ST function to financial data, finding evidence of sign asymmetry in the mean.

We next consider a specification of the second-order logistic function in van Dijk, Teräsvirta, and Franses (2002).

$$F(z_{t-d};\gamma,\boldsymbol{c}) = \frac{1}{1+\exp\left\{\frac{-\gamma(z_{t-d}-c_1)(z_{t-d}-c_2)}{s_z}\right\}}, \quad c_1 < c_2, \tag{3}$$

where now $\boldsymbol{c} = (c_1;c_2)'$, as proposed by Jansen and Teräsvirta (1996). Figure 1 shows some examples for the second-order ST function for various values of the smoothness parameter γ when $c_1 = -1.5$ and $c_2 = 1.5$. We observe that smaller values of γ cause smoother, slower transitions, while $\gamma \geq 20$ is effectively a sharp or abrupt transition. When $\gamma = 20$, the transition function starts at 1, then decreases to zero during the range of (c_1, c_2), and then increases back to one again.

When $\gamma \to 0$, both logistic functions become equal to a constant (equal to 0.5), and when $\gamma = 0$, the ST-GARCH model reduces to a linear GARCH model. If $\gamma \to \infty$, the model with the second-order logistic function becomes linear, whereas if $\gamma \to \infty$ and $c_1 \neq c_2$, the function $F(z_{t-d}; \gamma, \boldsymbol{c})$ is equal to 1 for $z_t < c_1$ and $z_t > c_2$, and equal to 0 in-between. Hence, the SETAR model with the particular transition function nests a restricted three-regime SETAR model, with a restriction of outer regimes being identical. For the third-type function, we consider the exponential function:

$$F(z_{t-d}; \gamma, c) = 1 - \exp\left\{ \frac{-\gamma(z_{t-d} - c)^2}{s_z} \right\}. \tag{4}$$

The behavior of y_t depends on the size of the deviation from z_t. The exponential ST is an even function, which captures size asymmetry, or asymmetric responses to the magnitude of $z_{t-d} - c$. See Granger and Teräsvirta (1993) and Teräsvirta (1994) for applications of STAR with the exponential ST (EST) model. Chen et al. (2010) apply STAR with the EST function to daily stock markets, finding evidence of size asymmetry in mean and volatility, while the most favored transition variable is the intra-day range. A limitation of the exponential function (4) is that for either $\gamma \to 0$ or $\gamma \to \infty$, the function collapses to a constant (equal to 0 and 1, respectively). Hence, the model becomes linear in both cases and the exponential STAR model does not nest a self-exciting TAR model as a special case (see van Dijk, Teräsvirta, and Franses 2002 for details).

We further allow the parameter γ varying in the ST-GARCH with (3) to incorporate a different effect (smooth transition function) for the mean and variance. While Gerlach and Chen (2008) illustrate sufficient restrictions for volatility equations, they only consider the first-order ST function. In this paper, we utilize the same restrictions upon the three ST-GARCH models. The non-negativeness of the conditional variance is:

$$\alpha_0^{(1)} > 0, \alpha_i^{(1)} > 0, \beta_i^{(1)} > 0 \;\; \sum_i \left(\alpha_i^{(1)} + \alpha_i^{(2)} \right) > 0 \;\; \sum_j \left(\beta_j^{(1)} + \beta_j^{(2)} \right) > 0. \tag{5}$$

The covariance-stationary restriction is as below:

$$\sum_i \left(\alpha_i^{(1)} + 0.5\alpha_i^{(2)} \right) + \sum_j \left(\beta_j^{(1)} + 0.5\beta_j^{(2)} \right) < 1. \tag{6}$$

These conditions can also be found in Anderson, Nam, and Vahid (1999). In order to allow for possible explosive volatility and to ensure a proper prior, Gerlach and Chen (2008) generalize the above two restrictions as follows:

$$\alpha_0^{(1)} < b_1, \beta_i^{(1)} < b_2, \ \sum_i \alpha_i^{(1)} + \sum_j \beta_j^{(1)} < b_3, \tag{7}$$

where b_1, b_2, and b_3 are user-specified. In this study, we let b_2, $b_3 \geq 1$ to allow explosive behavior.

3 Bayesian inference

In this section, we use the Bayesian approach to carry out our parameter estimations. In the literature, many papers have utilized the ordinary least squares method and the nonlinear least squares method to estimate parameters. Let $\boldsymbol{\theta} = (\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{c}, \boldsymbol{\gamma}, \nu, d)'$, where $\boldsymbol{\phi}_j$ =$(\phi_0^{(j)}, \ldots$ and $\phi_p^{(j)})'$, $\boldsymbol{\alpha}_j$ = $(\alpha_0^{(j)}, \ldots, \alpha_g^{(j)}, \beta_1^{(j)}, \ldots, \beta_k^{(j)})'$. Note that $\boldsymbol{c} = (c_1, c_2)'$ and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)'$ if the ST function belongs to the second-order ST-GARCH models, otherwise, $\boldsymbol{c} = c$ and $\gamma_1 = \gamma_2$. We allow for a higher degree of flexibility in model parameters for this model. The notation $\boldsymbol{y}^{1,n}$ denotes $(y_1, \ldots, y_n)$. The conditional likelihood function for the double ST model is:

$$L(\boldsymbol{\theta} \mid \boldsymbol{y}^{s+1,n}) = \prod_{t=s+1}^{n} \left\{ \frac{1}{\sqrt{h_t}} p_\epsilon \left(\frac{y_t - \mu_t}{\sqrt{h_t}} \right) \right\},$$

where p_ϵ is the density function for ϵ_t, n is the sample size, $s = \max\{p, g, q, d\}$, $h_t = \mathrm{Var}(y_t|\mathcal{F}_{t-1})$, and $\mu_t = E(y_t|\mathcal{F}_{t-1})$, with $\mathcal{F}_{t-1}$ being the information set. Based on the empirical evidence, the empirical density function has a higher peak and longer tails than the normal density. This phenomenon is common for daily stock returns. We consider $D(0,1) = t^*_\nu$ to be a standardized Student's t distribution, which captures the conditional leptokurtosis observed in financial return data. The conditional likelihood function becomes:

$$L(\boldsymbol{\theta} \mid \boldsymbol{y}^{s+1,n}) = \prod_{t=s+1}^{n} \left\{ \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{(\nu-2)\pi}} \frac{1}{\sqrt{h_t}} \left[1 + \frac{(y_t - \mu_t)^2}{(\nu-2)h_t} \right]^{-\frac{\nu+1}{2}} \right\}.$$

Aside from fat tails, empirical distributions of asset returns may also be skewed. To handle this additional characteristic of asset returns, the Student's t distribution has been modified to become a skew Student's t distribution. There are many versions of skew Student's t distribution, but we adopt the approach of Hansen (1994), which has zero mean and unit variance. The probability density function of skew Student's t defined by Hansen (1994) is as follows:

$$p_\epsilon(\epsilon_t|\nu, \eta) = \begin{cases} bc\left[1 + \frac{1}{\nu-2}\left(\frac{b\epsilon_t + a}{1-\eta}\right)^2\right]^{-(\nu+1)/2} & \text{if} \quad \epsilon_t < -\frac{a}{b} \\ bc\left[1 + \frac{1}{\nu-2}\left(\frac{b\epsilon_t + a}{1+\eta}\right)^2\right]^{-(\nu+1)/2} & \text{if} \quad \epsilon_t \geq -\frac{a}{b} \end{cases}, \tag{8}$$

where degrees of freedom ν and skewness parameter η satisfy $2 < \nu < \infty$, and $-1 < \eta < 1$, respectively. The constants a, b, and c are fixed as:
$a = 4\eta c\left(\frac{\nu-2}{\nu-1}\right), \quad b^2 = 1 + 3\eta^2 - a^2, \quad c = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi(\nu-2)}}.$

This distribution already has zero mean and unit variance. We use the notation $\mathrm{St}(\nu, \eta)$. The standardized Student's t distribution is a special case of this skew Student's t distribution, when $\eta = 0$.

3.1 The setup priors

Bayesian inference requires specifying a prior distribution for the unknown parameters, combined with the likelihood function. We assume the parameters $(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{c}, \boldsymbol{\gamma}, \nu, d)$, are a priori independent. An estimation of the smoothing parameter and its identification as it tends to zero have proven a challenge for both classical and Bayesian approaches, because the likelihood function is not integrable for this parameter in an ST-GARCH model. To alleviate this identifiability problem as the speed of the transition parameter tends to zero, we adopt a specific prior formulation for the mean in Equation (1), based on George and McCulloch (1993) and extended by Gerlach and Chen (2008). Note that Gerlach and Chen (2008) only handle the ST-GARCH model with the first-order logistic function. We define the latent variable $\delta_j^{(i)}$, which determines the prior distribution of $\phi_j^{(i)}$, via a mixture of two normals:

$$\phi_j^{(i)}|\delta_j^{(i)} \sim (1-\delta_j^{(i)})N(0, k^2{\tau_j^{(i)}}^2) + \delta_j^{(i)}N(0, {\tau_j^{(i)}}^2), \quad j = 1, ..., p$$

$$\delta_j^{(i)}|\gamma = \begin{cases} 1, & \text{if } i = 1 \text{ or } \gamma > \xi \\ 0, & \text{if } i = 2 \text{ and } \gamma \leq \xi, \end{cases} \tag{9}$$

where $i = 1, 2$ denotes the regime and $j = 1, \ldots, p$ denotes the lag order of the AR mean terms in $\boldsymbol{\phi_j}$. Here, ξ is a specified threshold and $\gamma \leq \xi$ indicates that $\mathrm{F}(z_{t-d}; \gamma, c) \to 0.5$; that is, an AR-GARCH model. As suggested in Gerlach and Chen (2008), we choose k to be a small positive value, so that if $\gamma \leq \xi$ and $\delta_j^{(2)} = 0$, then the posterior value for the parameters $\phi_j^{(2)}$ will be weighted by the prior value towards 0.

A constrained uniform prior is taken for $p(\boldsymbol{\alpha})$, the constraint defined by the indicator $I(S)$, where S defines the constraints in Equations (5), (6), and (7). For γ, we choose the log-normal distribution, $\gamma \sim LN(\mu_\gamma, \sigma_\gamma^2)$. The prior for the delay lag, d, is a discrete, uniform variable:

$$P_r(d) = \frac{1}{d_0},$$

where $d = 1, \ldots, d_0$. For ν degrees of freedom, we define $\rho = \frac{1}{\nu}$ and set it to $I(\rho \in [0, 0.25])$ (see Chen, Chiang, & So, 2003, for more details). We choose the flat priors for the threshold parameters in three ST functions, which are described as follows.

The first-order and the exponential ST function:

When we consider a double ST model with one threshold value, a flat prior on the threshold limit c is Unif(b_{z1}, b_{z2}), where (b_{z1}, b_{z2}) are chosen as suitable percentiles of z to allow a reasonable sample size in each regime for inference.

The second-order ST function:

Two threshold values in the second-order ST function are much more complicated and need to be constrained in two ways: the first ensures that $c_1 < c_2$ as required, while the second ensures that a sufficient sample size exists in each regime for estimation. For this second constraint, a set of ranges can be set, as relevant percentiles of the sample size n, to ensure that at least $100h$ $(0 < h < 1)$ percent of the observations are contained in each regime, as suggested by Chen, Gerlach, and Lin (2010) and Chen, Gerlach, and Liu (2011). The general priors for c_1 and c_2 are:

$$c_1 \sim \text{Unif}(lb_1, ub_1);$$

$$c_2 | c_1 \sim \text{Unif}(lb_2, ub_2),$$

where lb_1 and ub_1 are the $\wp_{h_1}$ and $\wp_{1-h_1-h_2}$ percentiles of z_t, respectively. For example, if $h_1 = h_2 = 0.1$, then $c_1 \in (\wp_{0.1}, \wp_{0.8})$. Furthermore, we set $ub_2 = \wp_{(1-h_2)}$ and $lb_2 = c_1 + c^*$, where c^* is a selected number that ensures $c_1 + c^* \leq c_2$ and at least $100h_2\%$ of observations are in the range (c_1, c_2).

3.2 Posteriors

The posteriors are proportional to the product of the likelihood function and the priors, or in other words:

$$p(\boldsymbol{\theta}_l \mid \boldsymbol{y}^{s+1,n}, \boldsymbol{\theta}_{\neq l}) \propto p(\boldsymbol{y}^{s+1,n} \mid \boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}_l \mid \boldsymbol{\theta}_{\neq l}),$$

where $\boldsymbol{\theta}_l$ is a parameter group, $p(\boldsymbol{\theta}_l)$ is its prior density, and $\boldsymbol{\theta}_{\neq l}$ is the vector of all model parameters, except for $\boldsymbol{\theta}_l$. The delay parameter d is obtained by sampling from the conditional multinomial distribution with posterior probabilities as follows:

$$p(d = j \mid \boldsymbol{y}^{s+1,n}, \boldsymbol{\theta}_{\neq d}) = \frac{p(\boldsymbol{y}^{s+1,n} \mid d = j, \boldsymbol{\theta}_{\neq d})}{\sum_{j=1}^{d_0} p(\boldsymbol{y}^{s+1,n} \mid d = j, \boldsymbol{\theta}_{\neq d})}, \quad j = 1, \ldots, d_0. \tag{10}$$

Since the posterior distributions for parameters $(\boldsymbol{\phi}_j, \boldsymbol{\alpha}_j, \nu, \gamma, \boldsymbol{c})$, with $j = 1, 2$ are not standard forms, we turn to the MCMC method. For our parameters $(\boldsymbol{\phi}_j, \nu, \gamma, \boldsymbol{c})$, and $j = 1, 2$, we estimate parameters by exercising the Metropolis-Hasting (MH) algorithm. For the GARCH parameter $\boldsymbol{\alpha}_j$, we apply a random walk MH algorithm before burn-in period and use the independent kernel MH (IK-MH) algorithm after the burn-in period, since IK-MH would speed up the convergence (see Gerlach & Chen, 2008, for more details). For details on the MCMC sampling scheme, random walk MH, and the IK-MH algorithm, please refer to Chen and So (2006).

4 Forecasting of VaR and volatility

Understanding volatility is vital for financial time series analysis. Predicting volatility is crucial for many functions in financial markets, such as estimation of VaR, options pricing, asset allocation, and many other applications. VaR is determined over a given time interval and could be exercised as a threshold value in order to avoid any downside risk on capital, given a specified probability α. Mathematically, VaR is defined as:

$$Pr(\triangle V(l) \leq -\text{VaR} \mid \mathcal{F}_{t-1}) = \alpha,$$

where $\triangle V(l)$ is an increment in the asset value over time period l, α is a given probability level, and $\mathcal{F}_{t-1}$ is as defined above.

A one-step-ahead VaR is the $\alpha\%$ quantile level of the conditional distribution $y_{n+1}|\mathcal{F}_n \sim D(\mu_{n+1}, h_{n+1})$, where h_{n+1} is given by one of the parametric models and D is the relevant error distribution. This predictive distribution is estimated via the MCMC simulation. The quantile VaR is then given by:

$$\text{VaR}_{n+1}^{[j]} = \mu_{n+1} - \left[D_{\alpha}^{-1}(\boldsymbol{\Theta}^{[j]}) \sqrt{h_{n+1}^{[j]}} \right], \tag{11}$$

where D^{-1} is the inverse CDF for the distribution D. For standardized Student's t errors,

$$D_{\alpha}^{-1} = \frac{t_{\alpha}(\nu^{[j]})}{\sqrt{\nu^{[j]}/(\nu^{[j]}-2)}},$$

where $t_{\alpha}(\nu^{[j]})$ is the αth quantile of a Student's t distribution with $\nu^{[j]}$ degrees of freedom, and $\nu^{[j]}$ is the jth iteration of ν. Hence, $\sqrt{\nu^{[j]}/(\nu^{[j]}-2)}$ is an adjustment term for a standardized Student's t with $\nu^{[j]}$ degrees of freedom. The final forecasted one-step-ahead VaR is the Monte Carlo posterior mean estimate:

$$\text{VaR}_{n+1} = \frac{1}{N-M} \sum_{j=M+1}^{N} \text{VaR}_{n+1}^{[j]}, \tag{12}$$

where N is the number of MCMC iterations, and M is the size of the burn-in sample.

4.1 VaR forecasting evaluation

In this section, we give the criteria for comparing and testing the VaR forecast models. The Basel Committee on Banking Supervision (established in 1996) proposed backtesting to evaluate the worst 1% expected loss over 250 trading days, so that at least one year of actual returns is compared with VaR forecasts. The common guides for comparing the performance are the number of violations ($I(y_t < -\text{VaR}_t)$) and the violation rate (VRate).

$$\text{VRate} = \frac{1}{m} \sum_{t=n+1}^{n+m} I\left(y_t < -\text{VaR}_t\right),$$

where n is the in-sample period size, and m is the forecast size. Naturally, a VRate close to nominal α is desirable. Furthermore, under the Basel Accord, models that over-estimate risk (VRate $< \alpha$) are preferable to those that under-estimate risk levels.

We are greatly interested in the magnitude of the VaR exceedance rather than simply whether or not an exceedance occurred. A backtest can be based on a function of the observed profit or loss and the corresponding model VaR. This would result in the construction of a general loss function, $L(\text{VaR}_t, y_t)$, which could be evaluated using past data on profits and losses and the reported VaR series. Lopez (1999a,b) suggests this approach to backtesting as an alternative to the approach that focuses exclusively on the hit series. We consider two loss functions that measure the difference between the observed loss and the VaR in cases where the loss exceeds the reported VaR measure.

$$\Psi_1(\text{VaR}_t, y_t) = \begin{cases} 1 + (y_t - (-\text{VaR}_t))^2 & \text{if } y_t < -\text{VaR}_t \\ 0 & \text{if } y_t \geq -\text{VaR}_t \end{cases}, \tag{13}$$

$$\Psi_2(\text{VaR}_t, y_t) = \begin{cases} 1 + \Big| y_t - (-\text{VaR}_t) \Big| & \text{if } y_t < -\text{VaR}_t \\ 0 & \text{if } y_t \geq -\text{VaR}_t \end{cases}. \tag{14}$$

When an exception takes place, risk model is to be penalized. Hence we prefer to have a lower average loss value (between two models), defined as the average of these penalty scores:

$$\mathbf{\Psi}_i = \frac{1}{m} \sum_{t=n+1}^{n+m} \Psi_i(\text{VaR}_t, y_t),\ i = 1, 2.$$

4.2 Backtesting methods

We further consider three backtesting methods for evaluating and testing the accuracy of VaR models. The unconditional coverage (UC) test of Kupiec (1995) - a likelihood ratio test that the true violation rate equals α; the conditional coverage (CC) test of Christoffersen (1998) - a joint test, combining a likelihood ratio test for independence of violations and the UC test; and the dynamic quantile (DQ) test of Engle and Manganelli (2004). The details of the processes are given below.

- **The UC test of Kupiec (1995):** As stated in Christoffersen (1998), the UC test looks at the unconditional probability of a violation that must be equal to the coverage rate α, with the LRT being:
$$LR_{uc} = 2\log\left[\frac{\hat{\alpha}^X(1-\hat{\alpha})^{m-X}}{\alpha^X(1-\alpha)^{m-X}}\right] \sim \chi_1^2,$$
where $X =$ number of violations, $m =$ total number of observations, and $\hat{\alpha} = X/m$.

- **The CC test of Christoffersen (1998):** The CC test is a joint test that combines a likelihood ratio test for independence of violations and the UC test, where the independence hypothesis stands for VaR violations observed at two different dates being independently distributed.
$$LR_{ind} = 2\log\left(\frac{L_1}{L_0}\right) \ ; \ LR_{ind} \sim \chi_1^2.$$
We define T_{ij} as the number of days when condition j occurred at present status, and assuming that condition i occurred on the previous day, we get:
$$i, j = \begin{cases} 1, & \text{if violation occurs} \\ 0, & \text{if no violation occurs,} \end{cases} \tag{15}$$
and $L_1 = \prod_{i=0}^{1}(1-\pi_{i1})^{T_{i0}}\pi_{i1}^{T_{i1}}, \ L_0 = (1-\pi)^{\sum_{i=0}^{1}T_{i0}}\pi^{\sum_{i=0}^{1}T_{i1}}$. $\pi_{i1} = T_{i1}/(T_{i0}+T_{i1})$, and $\pi = (T_{01}+T_{11})/m$, with m being the total number of observations. Thus, the joint CC test is a chi-square test, in which $LR_{cc} = LR_{uc} + LR_{ind}$, when $LR_{cc} \sim \chi_2^2$.

- **The DQ test of Engle and Manganelli (2004):** The DQ test is based on a linear regression model of the hits variable on a set of explanatory variables including a constant, the lagged values of the hit variable, and any function of the past information set suspected of being informative. H_0: $H_t = I(y_t <$ -VaR$_t) - \alpha$ is independent of $\boldsymbol{W}$. The test statistic is:
$$\text{DQ}(q) = \frac{\boldsymbol{H}'\boldsymbol{W}\left(\boldsymbol{W}'\boldsymbol{W}\right)^{-1}\boldsymbol{W}'\boldsymbol{H}}{\alpha(1-\alpha)},$$
where $\boldsymbol{W} =$ lagged observations, hits, etc. and $q = 4$. This is the same setting as in Engle and Manganelli (2004); DQ$(q) \sim \chi_q^2$.

4.3 Volatility proxies

Though volatility is unobservable, we consider the following three proxy variables. One is based on absolute returns and the two others are range-based proxies, using root mean square error (RMSE), mean absolute deviation (MAD), and quasi Log-likelihood (QLIKE) as loss functions. Patton and Sheppard (2009) show that the QLIKE loss function is more robust to noise in the volatility proxy. The first range-based proxy is like that of Parkinson (1980), the second is based on Alizadeh, Brandt, and Diebold (2002) and employed by Lin, Chen, and Gerlach (2012). The formulae for the three volatility proxies are the following.

1. Proxy 1: $|y_t|$.

2. Proxy 2: $R_t/\sqrt{4\ln(2)}$; (Parkinson, 1980).

3. Proxy 3: $\exp[\ln(R_t) - 0.43 + 0.29^2/2]$; (Alizadeh, Brandt, and Diebold, 2002),

where $R_t = (\max P_t - \min P_t) \times 100$, P_t is the log price index at time t, and y_t is the ith intra-day return on day t. We denote $\hat{h}_{i,t+1|t}$, $i = 1, \ldots, 3$ as the ith competitive one-step-ahead forecast of h_{t+1}, where m is the forecast period size. Three measures are calculated as follows:

$$
\begin{aligned}
\text{RMSE} &= \left[\frac{1}{m}\sum_{t=n}^{n+m-1}(e_{i,t+1|t})^2\right]^{0.5}, \\
\text{MAD} &= \frac{1}{m}\sum_{t=n}^{n+m-1}\left|e_{i,t+1|t}\right|, \\
\text{QLIKE} &= \frac{1}{m}\sum_{t=n}^{n+m-1}\tilde{e}_{i,t+1|t} - \log\left(\tilde{e}_{i,t+1|t}\right) - 1, \\
\text{where} &\quad e_{i,t+1|t} = \sqrt{h_{t+1}} - \sqrt{\hat{h}_{i,t+1|t}}, \quad \tilde{e}_{i,t+1|t} = \frac{\sqrt{h_{t+1}}}{\sqrt{\hat{h}_{i,t+1|t}}} \quad i = 1, \ldots, 3.
\end{aligned}
$$

When comparing the models, it is favorable to have smaller error values under the three criteria.

5 Simulation study

We perform simulation studies for the Bayesian estimation to examine the effectiveness of the MCMC sampling scheme. Considering finite sample properties and consistencies of the MCMC estimators, 500 replications are generated, with sample size $n = 2000$. We consider the second-order

ST-GARCH model with a skew Student's t distribution as follows:

$$
\begin{aligned}
y_t &= (0.1 + 0.4y_{t-1}) + F_1(z_{t-1})(0.1 - 0.25y_{t-1}) + a_t, \\
a_t &= \sqrt{h_t}\epsilon_t, \qquad \epsilon_t \overset{\text{i.i.d.}}{\sim} SK(7, -0.4) \\
h_t &= (0.15 + 0.2a_{t-1}^2 + 0.7h_{t-1}) + F_2(z_{t-1})(-0.1 - 0.1a_{t-1}^2 - 0.2h_{t-1}), \\
F_i(z_{t-1}) &= \frac{1}{1 + \exp\left\{\frac{-\gamma_i(z_{t-1}-(-0.35))(z_{t-1}-0.3)}{s_z}\right\}},
\end{aligned} \tag{16}
$$

where $(\gamma_1, \gamma_2) = (4, 10)$, z_t is the daily returns of the S&P500 index and s_z is its sample standard deviation. Since under the proposed model, it would be difficult to generate a series of z_t, we use S&P500 returns instead. Orders p, g, and q are all set to 1. The maximum delay, d_0, is chosen to be 3. The initial values for each parameter are $\boldsymbol{\phi}_1 = (0, 0)$, $\boldsymbol{\phi}_2 = (0, 0)$, $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = (0.01, 0.1, 0.1)$, $\nu = 100$, $\gamma = 30$, and $(c_1, c_2) = (0, 0.1)$.

Based on Figure 1, $F(.)$ becomes a sharp or abrupt transition when $\gamma > 20$. Therefore, we consider only choices of $(\mu_\gamma, \sigma_\gamma)$ that ensure the prior density becomes small for $\gamma > 20$. We establish two setups of prior information for $\gamma_i, i = 1, 2$, $LN(\mu_\gamma, \sigma_\gamma) = (1.609, 0.767)$ and $(1.609, 0.575)$, in which the densities are in Figure 2. We set the hyper-parameters to $(\xi, k) = (0.5, 0.001)$ in the mixture specification (9), with $\tau_i = 0.35$, c_1 $\sim$Unif$(\wp_{0.2}, \wp_{0.7})$, and $c_2|c_1$ $\sim$Unif$(c_1 + c^*, \wp_{0.8})$. Hence, c^* is chosen, which leads to at least 10% of observations in-between. We also set $b_1 = s_y^2$, $b_2 = 1$, and $b_3 = 1.1$ in (7) allowing for possible explosiveness in the variance equation. Note that these settings are suggested by Gerlach and Chen (2008).

We use a burn-in sample of $M = 10,000$ and a total sample of $N = 30,000$ iterations, but only every 2nd iteration in the sample period for inference. Parameter estimation results are in Table 1, in which numbers are the averages of posterior mean, median, standard deviation, and 95$\%$ credible interval for 500 replications, except for the parameter d, which is the delay lag. We provide the average, median, and standard deviation among 500 posterior modes in the last row in Table 1. They 100% correctly indicate d=1 for each 500 replications.

We extensively examine trace plots and the autocorrelation function (ACF) plots to confirm convergence and to infer adequate coverage. We observe the trace plots for parameters that converge immediately. The ACF plots cut off fairly quickly, which means that the MCMC mixing is fast and the autocorrelation is low. Those plots are not shown to save space.

We assume the same prior density for γ_1 and γ_2 - that is, we do not have any restriction about the magnitude of γ_1 and γ_2. The estimates of $(\gamma_1,\gamma_2) = (4, 10)$ are $(5.04, 10.03)$ and $(4.95, 8.68)$ based on priors 1 and 2, which are sound. The average standard deviations for prior 2 are slightly smaller versus those of prior 1. These simulation results indicate that the posterior estimates obtained by the proposed sampling scheme are reliable. Non-Bayesian methods are unable to accomplish this desired purpose for estimation.

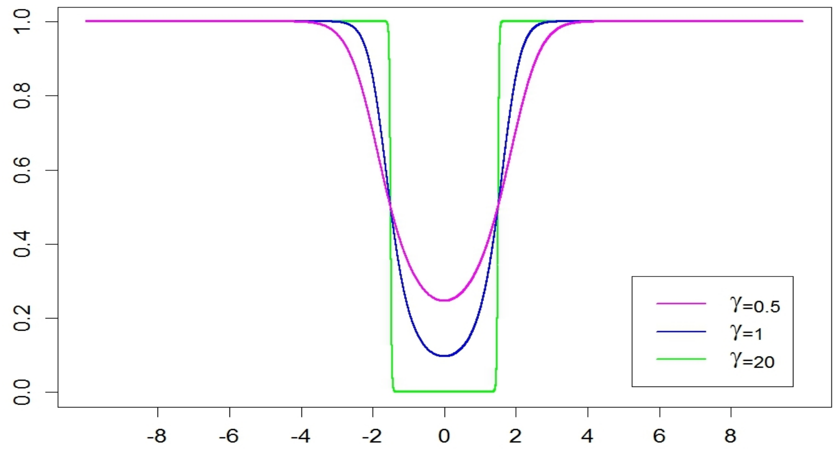

Figure 1: Plots of the second-order ST function for c_1=-1.5 and c_2=1.5.

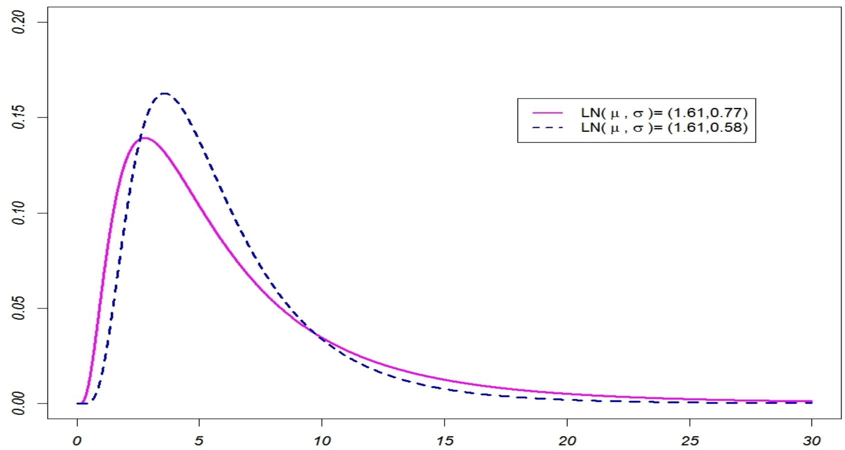

Figure 2: Two prior densities for γ.

6 Empirical study

To examine the performance of the models under highly varied market conditions, this study examines two distinct forecasting periods. The first complete dataset is divided into two: an in-sample period from January 1, 2004 to October 31, 2011, and 500 out-of-sample forecast days, from November 1, 2011

Table 1: Simulation results for the second-order ST-GARCH model in (16) based on n=2000 and obtained from 500 replications.

	True	Mean	Med	Std	Mean	Med	Std
			Prior 1			Prior 2	
$\phi_0^{(1)}$	-0.10	-0.0992	-0.0991	0.0345	-0.1012	-0.1009	0.0347
$\phi_1^{(1)}$	0.30	0.3072	0.3048	0.0833	0.3019	0.2997	0.0834
$\phi_0^{(2)}$	0.10	0.0980	0.0966	0.0412	0.1026	0.1021	0.0412
$\phi_1^{(2)}$	-0.25	-0.2605	-0.2579	0.0993	-0.2467	-0.2472	0.0991
$\alpha_0^{(1)}$	0.15	0.1629	0.1672	0.0295	0.1632	0.1651	0.0289
$\alpha_1^{(1)}$	0.20	0.2080	0.2034	0.0762	0.2170	0.2153	0.0768
$\beta_1^{(1)}$	0.70	0.6503	0.6524	0.1141	0.6651	0.6678	0.1116
$\alpha_0^{(2)}$	-0.10	-0.1084	-0.1110	0.0330	-0.1089	-0.1106	0.0320
$\alpha_1^{(2)}$	-0.10	-0.1083	-0.1001	0.0845	-0.1140	-0.1109	0.0858
$\beta_1^{(2)}$	-0.20	-0.1819	-0.1859	0.1394	-0.1905	-0.1933	0.1349
ν	7.00	7.1252	6.9782	1.0762	7.0909	6.9636	1.0733
η	-0.40	-0.4016	-0.4018	0.0292	-0.3963	-0.3966	0.0291
γ_1	4.00	5.0494	4.3081	3.1122	4.9478	4.4164	2.3318
γ_2	10.00	10.0281	9.0847	4.5053	8.6667	8.1056	3.1059
c_1	-0.35	-0.3356	-0.3326	0.1077	-0.3333	-0.3303	0.1076
c_2	0.30	0.2886	0.2866	0.1075	0.2849	0.2815	0.1068
d^*	1.00	1	1	0	1	1	0

Average of posterior modes and median of posterior modes for d.

to late October or mid-November 2013. Small differences in end-dates across markets occurred due to different market trading days. This time frame is a period after the effects of the global financial crisis hit the markets. The study includes ten European stock markets, seven Asian stock markets, and 1 North American market and 1 South American market, making 19 stock markets in all. We utilize three regions for the daily closing prices of stock markets, including (i) Americas: the S&P500 (U.S.) and the Bovespa Index (Brazil). (ii) Asia: KOSPI (South Korea), HANG SENG Index (Hong Kong), Nikkei 225 (Japan), CNX 500 (India), SHANGHAI SE A SHARE (China), TAIEX (Taiwan), and SET Index (Thailand). (iii) Europe: FTSE 100 (U.K.), DAX 30 (Germany), CAC 40 (France), AEX Index (Netherlands), PSI 20 (Portugal), MIB Index (Italy), ISEQ (Ireland), Athex Composite Index (Greece), RTS Index (Russia), and IBEX 35 (Spain).

To examine how the models perform during the recent financial crisis period (2007-2009) and evaluate how the crisis affects risk management, a second time span in considered: a learning period from January 4, 2000 to December 31, 2006 and a second validation or out-of-sample forecast evaluation window: January 3, 2007 to December 30, 2009. We mainly focus on S&P500 and Nikkei 225 for this financial turmoil period.

All data are obtained from Datastream International. The returns are the difference of the logarithm of the daily price index:

$$r_t = [\ln(P_t) - \ln(P_{t-1})] \times 100,$$

where P_t is the closing index on day t. Table 2 shows summary statistics for the 19 markets during the in-sample period from January 1, 2004 to October 31, 2011. The statistics include stock-index return means, extreme values, standard deviations, skewness, kurtosis, the Jarque-Bera normality test, and Liung-Box Q(5) values for both returns and squared returns. As per the characteristics of financial data, the daily return has heavy tails and is negatively skewed (exceptions are for Hong Kong, Germany, France, Greece, and Spain). The normality test exhibits a clear rejection for each market by the Jarque-Bera normality test under a 1% significant level. Furthermore, high volatility is clearly evident during the global financial crisis period (around late - 2008 to 2009).

The DT-GARCH model is a special case of the first-order ST-GARCH (1STF-GARCH) model when the smoothness parameter γ goes to infinity (see Chen and So, 2006). We state the DT-GARCH(1,1) model as follows:

$$y_t = \begin{cases} \phi_0^{(1)} + \phi_1^{(1)} y_{t-1} + a_t & y_{t-d} < c \\ \phi_0^{(2)} + \phi_1^{(2)} y_{t-1} + a_t & y_{t-d} \geq c \end{cases}, \tag{17}$$

$$a_t = \sqrt{h_t}\epsilon_t, \qquad \epsilon_t \overset{\text{i.i.d.}}{\sim} t_\nu^*,$$

$$h_t = \begin{cases} \alpha_0^{(1)} + \alpha_1^{(1)} a_{t-1}^2 + \beta_1^{(1)} h_{t-1} & y_{t-d} < c \\ \alpha_0^{(2)} + \alpha_1^{(2)} a_{t-1}^2 + \beta_1^{(2)} h_{t-1} & y_{t-d} \geq c \end{cases}, \tag{18}$$

where d is the delay lag and c is the threshold value. Parameters in the DT-GARCH model, $\phi_0^{(j)}$, $\phi_1^{(j)}$, $\alpha_0^{(j)}$, $\alpha_1^{(j)}$, $\beta_1^{(j)}$, (where $j = 1, 2,$) c, d, and ν, are estimated by the Bayesian method proposed by Chen and So (2006). When $\alpha_0^{(2)} = \beta_1^{(2)} = 0$, $d = 1$, and $c = 0$ in Equation (18), then this model becomes a special case: GJR-GARCH model (Glosten, Jaganathan, Runkle, 1993).

We report Bayesian estimates for the U.S. market during the in-sample period (from January 1, 2004 to October 31, 2011) based on the five nonlinear heteroskedastic models: first-order logistic function (1STF-GARCH) with Student's t error, second-order logistic function ST-GARCH (2STF-GARCH) with Student's t and skew Student's t errors, the exponential function ST-GARCH (ESTF-GARCH), and the DT-GARCH model with Student's t errors. The priors' settings for ST-GARCH are the same as in the simulation study, and the estimation is based on a total of 30,000 MCMC iterations, discarding the first 10,000 iterations as a burn-in period. Those estimations would perform differently when the rolling window span is allowed to move. Table 3 presents the estimated posterior median and standard deviation of parameters. To save space, we do not report 95% credible intervals here.

The majority of coefficients in mean equations are insignificant, which are indicated by the 95% credible intervals. Allowing AR(1) in the conditional mean helps account for possible asymmetric autocorrelations in the returns. The delay lag d is not always fixed and swings from 1 to 3, based on learning periods. In order to show further justification about the same effect or the same smooth transition function in the mean and variance, we allow Model "2STsk-GARCH" to incorporate different smooth transition functions for the mean and variance and skew Student's t errors. The estimated skew parameter and degrees of freedom are -0.149 and 6.955, indicating the skew Student's t assumption is appropriate. However, the estimates of γ_1 and γ_2 are 5.562 and 4.545, respectively. It seems that the same effect in the mean and variance due to the two estimated smoothness parameters is indistinguishable. In order to further examine about whether we should include different effects in mean and variance, we plan to use these five models for VaR forecasting.

In the out-of-sample period, a rolling window approach is used to produce a one-step-ahead forecasting of h_{n+1}, 1% VaR, and 5% VaR under the following 2 HS methods and 12 risk models: the HS-Short term with 25 observation days (ST), the HS-Long term with 100 observation days (LT), RiskMetrics (RM), AR(1)-GARCH(1,1) and AR(1)-GJR-GARCH(1,1) with three error distributions, 1STF-GARCH with Student's t error, second-order logistic function ST-GARCH (2STF-GARCH) with Student's t and skew Student's t errors, the exponential ST-GARCH (ESTF-GARCH), and the DT-GARCH model with Student's t errors. For the global financial crisis period, realized volatility (RV) models with three error probability distributions are considered for U.S. and Japan stock markets.

Post-global financial crisis period

Tables 4 and 5 report the empirical results for the 19 markets: VRate$/\alpha = 1$ under the 1% and 5% confidence levels based on a forecast period of 500 trading days. The ratio VRate$/\alpha = 1$ indicates a good VaR method/model. Figure 3 displays boxplots for VaR prediction performance over the 19 markets and 2 HS methods and 12 risk models at 1% and 5% levels. The figure illustrates that models with Gaussian errors and HS methods underestimate risk level at the 1% in all or most markets. Among these models, top three are G-sk, GJRsk, and 2STsk-GARCH models, in which the means of VRate are closest to nominal at the 1% level. These three best models have skew Student's t errors; clearly fat tails with additional skew characteristic are required in this forecasting time period. The results are different from $\alpha = 1\%$ to 5%. There are several models whose violation rates equal to or less than one at the 5% level.

We notice that the range of the other VRate$/\alpha$ varies from $[1.2, 5.2]$ for the Greece stock market at the 1% level, which is far away from 1. The worst performance for most of the risk models occurs in the Greece market. Three models - GARCH, GJR-GARCH, and 2STF-GARCH models with skew Student's t - along with 2STF-GARCH with Student's t errors stand out as performing the best across the European region at the $\alpha = 1\%$ level. The 2STF-GARCH model with Student's t error is favored by Ireland and Spain markets. We would like to point out that the banking crisis in Ireland in November 2010 further dented confidence in an already uncertain global financial market. It is estimated that Ireland owes well over $130 billion to German and British banks. The wide exposure of the crisis to the rest of the European market will likely weaken market confidence in that region in the coming months. Under these circumstances, the 2STF-GARCH model is a good choice for these markets during the post-global financial crisis period at a 1% level.

These results are also confirmed by Table 6. Based on the idea of Lopez (1993 a, b), we construct the following loss function in Table 6 to evaluate the performance of risk models (methods) that penalize those violation rates exceeding the α level by using the squared difference.

$$\Psi(\text{VRate}) = \begin{cases} \alpha + (\text{VRate} - \alpha)^2 & \text{if VRate} > \alpha \\ 0 & \text{if VRate} \leq \alpha \end{cases}. \tag{19}$$

Three models, GARCH, GJR-GARCH, and 2STF-GARCH models with skew Student's t perform the best at both 1% and 5% levels which yield the three lowest average loss values based on (19).

To evaluate the efficiency of risk measurement, Table 7 presents the results of Lopez's loss functions. Three models with skew Student's t errors, G-sk, GJRsk, and 2STsk-GARCH models, are the best three based on the median of quadratic loss and the absolute loss at the 1% level. It is clear that the choice of error distribution is highly important during this period. The best models turn out to be 2STsk-GARCH, GJR-n, and GJR-sk over quadratic loss and absolute loss at the 5% level. The family of GJR models is a good choice when we only consider risk at the 5% level.

Table 8 briefly describes the number of rejections for each model, over the 19 markets, at the 5% significance level, for each of the three tests considered: the UC, CC, and the DQ tests. Four lags are used, as stated in Engle and Manganelli (2004) for the DQ test. The "Total" states are the number of markets rejected by any backtests under each model. Under the ST method, all markets fail under the three backtests. At $\alpha = 1\%$, the ST method, the LT method, and the RM model are rejected in most of the markets, mainly by the DQ test. The 1STF-GARCH model and the DT-GARCH model have fewer rejections among tests in all the markets. At $\alpha = 5\%$, there are 10 models with only 0 or 1 rejections across markets.

Global financial crisis period

We next consider the recent financial crisis period (2007-2009) as an out-of-sample forecast evaluation window. We also investigate the realized GARCH models proposed by Hansen, Huwan and Shek (2012), where the daily returns and realized measure of volatility calculated using the intraday returns are jointly modelled. The realized measure of volatility calculated using the intraday returns may be subject to the bias caused by microstructure noise and non-trading hours. The realized GARCH model can adjust the bias in the realized measure. As a realized measure of volatility, we use the realized kernel calculated by taking account of the bias caused by microstructure noise (Barndorff-Nielsen et al. 2008). The realized kernel of the S&P 500 index is downloaded from the Oxford-Man Institute Realized Library (Heber et al., 2009) and that of the Nikkei 225 index is calculated using one-minute returns of the Nikkei 225 index obtained from the Nikkei NEEDS-tick data (Ubukata and Watanabe, 2014). We describe the realized GARCH model with three error distributions, given by the following three equations:

$$r_t = \sigma_t \zeta_t, \quad \zeta_t \overset{iid}{\sim} D(0,1) \tag{20}$$

$$\ln \sigma_t^2 = \omega + \beta \ln \sigma_{t-1}^2 + \gamma \ln x_{t-1} \tag{21}$$

$$\ln x_t = \xi + \varphi \ln \sigma_{t-1}^2 + \tau(\zeta_t) + u_t, \tag{22}$$

where r_t is the return and x_t is the realized kernel. Here, $\zeta_t \overset{iid}{\sim} D(0,1)$, $D(0,1)$ indicates a distribution that has mean 0 and variance 1, $u_t \overset{iid}{\sim} N(0, \sigma_u^2)$, and σ_t^2=var$(r_t|\mathcal{F}_{t-1})$ with $\mathcal{F}_t = \sigma(r_t, x_t, r_{t-1}, x_{t-1}, \ldots)$. Equation (22) is called a measurement equation, which relates the realized measure of volatility to the true volatility. If the realized measure were an unbiased estimator of the true volatility, then ξ and φ would be 0 and 1, respectively. Realized volatility, however, has a bias caused by microstructure noise and non-trading hours. Since we use the realized kernel, the bias caused by microstructure noise may be negligible. New York Stock Exchange and Tokyo Stock Exchange are open only for 6.5 hours and 5 hours, respectively, within a normal trading day and our realized kernels are calculated using the intraday returns only when the market is open. Thus, we should expect $\xi < 0$ or $\varphi < 1$.

In Equation (22), $\tau(\zeta_t) = \tau_1\zeta_t + \tau_2(\zeta_t^2 - 1)$ is utilized to generate an asymmetric response in volatility to return shocks. Three error distributions are used for the i.i.d. disturbances in each RV-type model in Equation (20). The choice D(0,1) is a standard Gaussian and labelled as RV-n. The Student's t (RV-t) and skew Student's t (RV-sk) distributions need to be standardized to have unit variance. We use the classical estimator, employing the "rugarch" package in R software, for modelling and forecasting RV models in (20)-(22) (see Ghalanos 2014). To save space, we do not provide the parameter estimation for RV models here, which are available from the authors upon request. However, we do observe that the estimated ξ is significantly negative and φ is significantly less than one for both U.S. and Japan stock markets. The estimate of τ_1 is significantly negative for both stock markets, indicating a negative correlation between today's return and tomorrow's volatility.

We construct the results of VRate/α and three backtests in Table 9. In the global financial crisis forecast period, however, all models significantly underestimate risk levels at the 1% and 5% quantiles and no model could be recommended as being accurate. All VRate/α values are greater than 1. For the Japan market, the violation rates of G-n, GJR-sk, 2ST-GARCH, and 2STsk-GARCH reach a minimum value (equal to 1.5) at the 1% level. For the U.S. market, RV-sk has a minimum VRate at the 1% level. Clearly, during the financial turmoil period, a skew error distribution with fat tails is very important to capture risk dynamics and level, at the 1% level, under a 1-day horizon. Most backtests for both markets are rejected at the 5% level. The measurements in Table 10 are based on Lopez's loss functions. It turns out that GJR-sk is the best model for the Japan market, while RV-sk and GJR-sk have the best performance for the U.S. market during financial turmoil periods.

Table 11 provides an evaluation of volatility forecasting based on three proxies and three loss functions. As the loss functions are judged under RMSE, MAD, and QLIKE, we prefer the model with the smallest value. The performances of volatility forecasts and VaR are in contrast to one another. Apparently, the DT-GARCH model and the ESTF-GARCH model are suitable under these criteria based on proxies (RM is better in some cases), but the VaR forecasts for them are not the best among the risk models. VaR estimates depend much more on the choice of distribution than volatility estimates do. However, when comparing the performance of the 2STF-GARCH model and the rest of models, the differences do not seem too large. We conclude that the 2STF-GARCH model is not excellent, but is still acceptable in volatility forecasting.

7 Conclusion

VaR is exercised as a threshold value in order to avoid downside risks on capital, given a specified probability. This paper evaluates performance of VaR forecasts across a range of competing parametric heteroskedastic models and non-parametric methods. Three variant ST functions are employed in order to capture the asymmetry in nonlinear, double threshold GARCH models. For a comparison, we also consider two popular asymmetric families: GJR-GARCH and DT-GARCHs models. Bayesian MCMC methods are employed on all heteroskedastic models (except the RV model) for estimation, inference, and forecasts. A simulation study shows that model parameters are well estimated for the 2STF-GARCH model with a different effect (smooth transition function) for the mean and variance and skew Student's

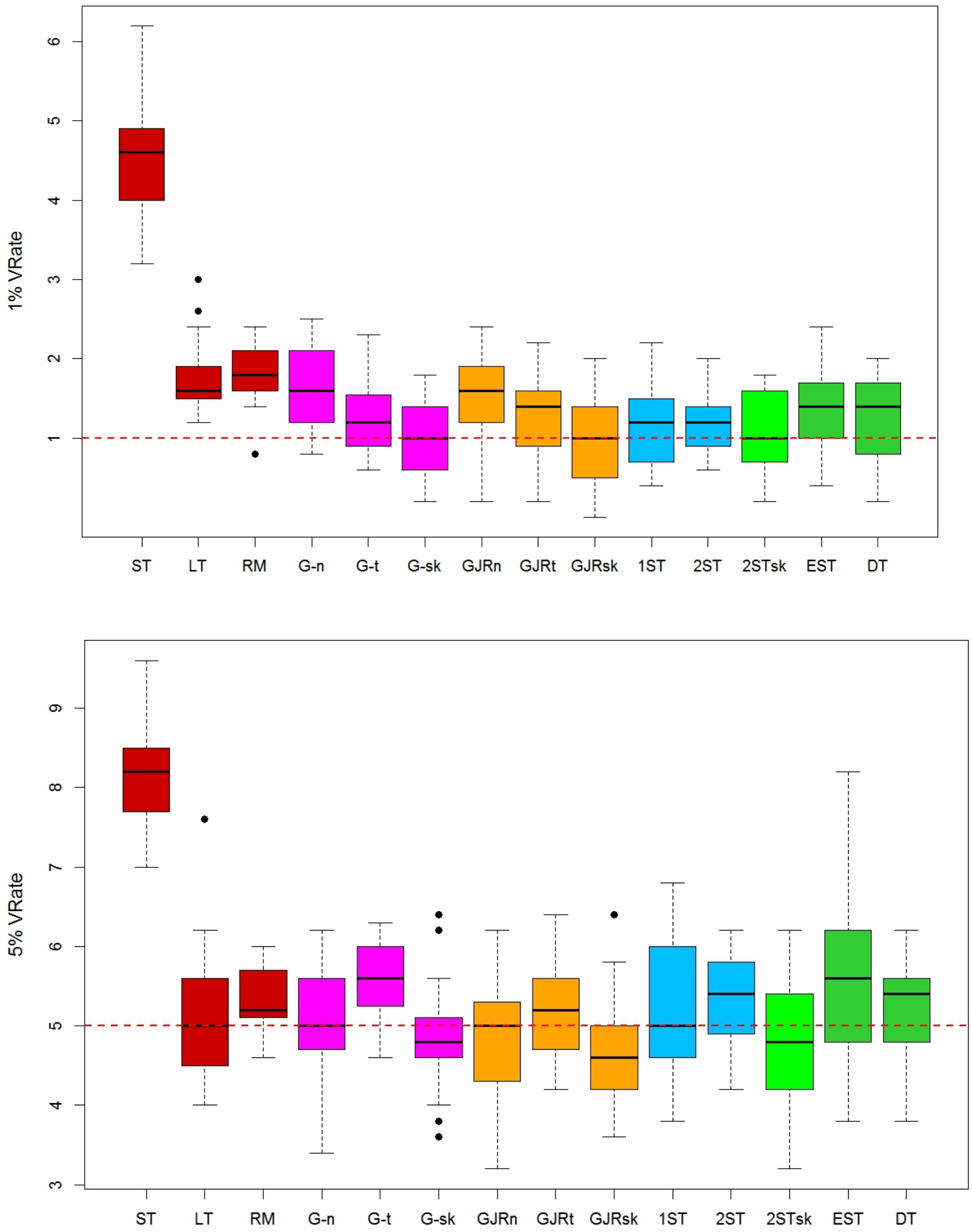

Figure 3: The boxplots for VaR prediction performance over the 19 markets and 2 HS methods and 12 risk models at the 1% and 5% levels.

t errors.

We evaluate two out-of-sample periods in light of the recent global financial crisis. For the post-global financial crisis, GARCH, GJR, and 2ST-GARCH with skew Student's t errors perform best at the 1% level based on Lopez's loss functions. The results show that the second-order logistic ST model is a good choice at the 1% level during the post-global financial crisis period, when compared to a range of existing alternatives. Both DT-GARCH and ESTF-GARCH models have a favorable out-of-sample volatility forecasting performance. Based on the empirical application, we conclude that higher moments (skewness and kurtosis) need to be explicitly modeled in order to obtain better VaR predictions.

For the global financial crisis period, all risk models underestimate the risk levels. We find that volatility asymmetry is most important for capturing risk, with skew errors also prominent, especially at the 1% level during the global financial period. In further works, we can also focus on expected shortfalls - that is, the expected number on the worst side, under a given percentage, which is more sensitive than VaR.

The use of our proposed Bayesian forecasting of nonlinear ST models to deal with some complex derivatives and to calculate their corresponding VaR formulae is of practical importance and theoretical interests. The Bayesian approach provides risk traders with the flexibility of adjusting their VaR models according to their subjective opinions. The findings of this research contribute to a better understanding of the performance of Bayesian forecasting of VaR based on various nonlinear ST models and hence could help securities traders or commercial banks in terms of valuating their risky portfolios.

Acknowledgements Cathy W.S. Chen's research is funded by the Ministry of Science and Technology, Taiwan (MOST 103-2118-M-035-002-MY2). Toshiaki Watanabe's financial supports are from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government through Grant-in-Aid for Scientific Research (No.25245034; 25245037; 26245028), Center for Financial Engineering Education at Hitotsubashi University and the Hitotsubashi Institute for Advanced Study.

References

Andersen, T.G., Bollerslev, T., Diebold, F.X., & Labys, P. (2001). The distribution of realized exchange rate volatility, *Journal of the American Statistical Association*, *96*, 42–55.

Alizadeh, S., Brandt, M.W., & Diebold, F.X. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance*, *57*, 1047-1092.

Anderson, H.M., Nam, K., & Vahid F. (1999). Asymmetric nonlinear smooth transition GARCH models. In P. Rothman (ed.) *Nonlinear Time Series Analysis of Economic and financial Data*, 191-207. Kluwer.

Basel Committee on Banking Supervision (1996). Supervisory framework for the use of 'backtesting' in conjunction with the internal models approach to market risk capital requirements. BIS: Basel.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde A., & Shephard, N. (2008). Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise", *Econometrica*, *76*, 1481-1536.

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

Chan, K.S., & Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, *7*, 179-190.

Chen, C.W.S., Chiang, T.C., & So, M.K.P. (2003). Asymmetrical reaction to US stock-return news: Evidence from major stock markets based on a double-threshold model. *Journal of Economics and Business*, *55*, 487-502.

Chen, C.W.S., Gerlach, R., Choy, S.T.B., & Lin, C. (2010). Estimation and inference for exponential smooth transition nonlinear volatility models. *Journal of Statistical Planning and Inference*, *140*, 719-733.

Chen, C.W.S., Gerlach, R., Hwang, B.B.K., & McAleer, M. (2012). Forecasting Value-at-Risk using nonlinear regression quantiles and the intra-day range. *International Journal of Forecasting*, *28*, 557-574.

Chen, C.W.S., Gerlach, R., & Lin, M.H. (2010). Falling and explosive, dormant, and rising markets via multiple-regime financial time series models. *Applied Stochastic Models in Business and Industry*, *26*, 28-49.

Chen, C.W.S., Gerlach, R., & Liu, F.C. (2011). Detection of structural breaks in a time-varying heteroskedastic regression model. *Journal of Statistical Planning and Inference*, *141*, 3367-3381.

Chen, C.W.S., & So, M.K.P. (2006). On a threshold heteroscedastic model.*Journal of Forecasting*, *22*, 73-89.

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

Engle, R.F. (1982). Autoregressive conditional heterosedasticity with estimates of variance of United Kingdom inflation. *Econometrica*, *50*, 987-1008.

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

George, E.I., & McCulloch, R.E. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association*, *88*, 881-889.

Gerlach, R., & Chen, C.W.S. (2008). Bayesian inference and model comparison for asymmetric smooth transition heteroskedastic models. *Statistics and Computing*, *18*, 391-408.

Gerlach, R., Chen, C.W.S., & Chan, N.C.Y. (2011). Bayesian time-varying quantile forecasting for value-at-risk in financial markets. *Journal of Business & Economic Statistics*, *29*, 481-492.

Ghalanos, H. (2014) *rugarch: Univariate GARCH models.*.R package version 1.3-1.

Glosten, L.R., Jagannathan, R., & Runkle, D.E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, *487*, 1779-1801.

Granger, C.W.J., & Teräsvirta, T. (1993). Modelling nonlinear economic relationships. *Oxford University Press, Oxford.*

Hansen, B. E. (1994) Autoregressive conditional density estimation. *International Economic Review*, *35*, 705-730.

Hansen, P. R., Huang, Z., & Shek, H. (2012) Realized GARCH: A Joint Model of Returns and Realized Measures of Volatility. *Journal of Applied Econometrics*, *27*, 877-906.

Heber, G., Lunde, A., Shephard, N., & Sheppard, K. (2009), Oxford-Man Institute's Realized Library, Version 0.1, Oxford-Man Institute, University of Oxford.

Jansen, E.S., & Teräsvirta, T. (1996). Testing parameter constancy and super exogeneity in econometric equations. *Oxford Bulletin of Economics and Statistics*, *58*, 735-768.

Jorion, O. (1997). Value at risk: The new benchmark for controlling market risk. Irwin Professional.

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

Lin, E.M.H., Chen, C.W.S., & Gerlach, R. (2012). Forecasting volatility with asymmetric smooth transition dynamic range models. *International Journal of Forecasting*, *28*, 384-399.

Lopez, J.A. (1999 a). Regulatory evaluation of Value-at-Risk models. *Journal of Risk*, *1*, 37-64.

Lopez, J.A.(1999 b). Methods for evaluating Value-at-Risk models. *Federal Reserve Bank of San Francisco Economic Review*, *2*, 3-17.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, *53*, 61-65.

Patton, A. J. & Shephard, K. (2009) Optimal combinations of realised volatility estimators. *International Journal of Forecasting*, *25*, 218-328.

Rabemananjara, R., & Zakoïan, J.M. (1993). Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics*, *8*, 31-49.

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, *89*, 208-218.

Teräsvirta, T., & Anderson, H. M. (1992). Characterizing nonlinearities in business cycles using smooth transition autoregressive models. *Journal of Applied Econometrics*, *7*, 119-136.

Ubukata, M. & Watanabe, T. (2014) Pricing Nikkei 225 options using realized volatility. *Japanese Economic Review*, *65*, 431-467.

van Dijk, D., Teräsvirta, T., & Franses, P.H. (2002). Smooth transition autoregressive models - a survey of recent developments. *Econometric Reviews*, *21*, 1-47.

Zakoïan, J.M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, *18*, 931-955.

Table 2: Summary statistics of market returns[a]: in-sample period from January 1, 2004 to October 31, 2011

	Mean	Min	Max	Std.	Skewness	Kurtosis	Jarque-Bera test	Q(5)	$Q^2(5)$
Americas									
Brazil	0.050	-12.100	13.680	1.9463	-0.081	8.263	< 0.001	0.024	< 0.001
U.S.	0.006	-9.470	10.960	1.3848	-0.312	13.161	< 0.001	< 0.001	< 0.001
Europe									
France	-0.005	-9.472	10.590	1.4814	0.065	10.559	< 0.001	< 0.001	< 0.001
Germany	0.022	-7.433	10.800	1.4383	0.052	10.354	< 0.001	0.025	< 0.001
Greece	-0.053	-10.210	13.430	1.7603	0.047	7.956	< 0.001	0.036	< 0.001
Ireland	-0.030	-13.960	9.733	1.6616	-0.575	9.966	< 0.001	0.536	< 0.001
Italy	-0.026	-8.598	10.880	1.5155	-0.043	10.172	< 0.001	0.001	< 0.001
Netherlands	-0.005	-9.590	10.030	1.4455	-0.174	11.957	< 0.001	< 0.001	< 0.001
Portugal	-0.007	-10.380	10.200	1.2099	-0.114	13.874	< 0.001	0.134	< 0.001
Russia	0.052	-21.200	20.200	2.3665	-0.454	14.085	< 0.001	< 0.001	< 0.001
Spain	0.007	-9.586	13.480	1.5068	0.150	11.685	< 0.001	0.001	< 0.001
U.K.	0.011	-9.266	9.384	1.2913	-0.150	11.307	< 0.001	< 0.001	< 0.001
Asia									
China	0.026	-9.261	9.033	1.8159	-0.294	5.920	< 0.001	0.078	< 0.001
Hong Kong	0.024	-13.580	13.410	1.7202	0.044	11.652	< 0.001	0.184	< 0.001
India	0.052	-12.880	15.030	1.7253	-0.488	11.131	< 0.001	< 0.001	< 0.001
Japan	-0.009	-12.110	13.230	1.6112	-0.569	12.236	< 0.001	0.230	< 0.001
South Korea	0.044	-11.170	11.280	1.5320	-0.578	9.212	< 0.001	0.233	< 0.001
Taiwan	0.013	-6.912	6.525	1.397	-0.432	5.996	< 0.001	0.0216	< 0.001
Thailand	0.012	-16.060	10.580	1.4913	-0.928	15.513	< 0.001	0.221	< 0.001

[a] The summary statistics results exclude the out-of-sample forecasting period.

Table 3: Bayesian estimation of parameters for 1STF, 2STF, ESTF-GARCH specifications, and DT-GARCH for the S&P500 index.

	1ST-GARCH		2ST-GARCH		2STsk-GARCH		EST-GARCH		DT-GARCH	
	Med	Std	Med	Std	Med	Std	Med	Std	Med	Std
$\phi_0^{(1)}$	0.533	0.302	0.029	0.055	-0.036	0.079	0.027	0.028	0.033	0.035
$\phi_1^{(1)}$	-0.177	0.162	-0.008	0.069	-0.150	0.103	-0.008	0.042	-0.084	0.040
$\phi_0^{(2)}$	-0.908	0.605	0.088	0.091	0.104	0.109	0.102	0.156	0.034	0.074
$\phi_1^{(2)}$	0.274	0.313	-0.087	0.102	0.099	0.132	-0.098	0.075	-0.033	0.064
$\alpha_0^{(1)}$	0.206	0.006	0.185	0.010	0.030	0.025	0.006	0.005	0.007	0.008
$\alpha_1^{(1)}$	0.345	0.013	0.273	0.008	0.128	0.039	0.006	0.006	0.133	0.018
$\beta_1^{(1)}$	0.354	0.007	0.331	0.012	0.681	0.057	0.885	0.018	0.947	0.019
$\alpha_0^{(2)}$	0.188	0.023	0.196	0.016	0.004	0.039	0.312	0.092	0.008	0.007
$\alpha_1^{(2)}$	0.167	0.017	0.133	0.016	-0.059	0.048	0.105	0.022	0.005	0.006
$\beta_1^{(2)}$	0.286	0.011	0.273	0.007	0.266	0.065	0.014	0.037	0.883	0.018
ν	4.081	0.112	4.174	0.231	6.955	1.298	6.784	1.194	7.220	1.309
η					-0.149	0.029				
γ	0.247	0.103	6.714	1.657			0.238	0.041	-	-
γ_1					5.562	3.424				
γ_2					4.545	2.867				
c	0.054	0.377	-	-			0.672	0.052	0.325	0.269
c_1	-	-	-0.326	0.186	0.183	0.131	-	-	-	-
c_2	-	-	0.568	0.123	0.518	0.094	-	-	-	-
d^*	1		1		2		1		1	

* denotes the posterior mode for d.

Table 4: VaR prediction performance using 2 HS methods and 12 risk models and 500 forecasted stock returns under $\alpha = 1\%$. VRate/α are given.

Markets	ST	LT	RM	G-n	G-t	G-sk	GJRn	GJRt	GJRsk	1STF GARCH	2STF GARCH	2STsk GARCH	ESTF GARCH	DT GARCH
Americas	(5.1)	(1.5)	(1.7)	(1.5)	(1.0)	(0.8)	(1.3)	(0.7)	(0.4)	(0.6)	(0.8)	(0.7)	(0.9)	(0.8)
Brazil	5.0	1.6	1.4	0.8	0.8	0.6	0.8	0.2	0.2	0.4	0.8	0.6	0.4	0.4
U.S.	5.2	1.4	2.0	2.2	1.2	1.0	1.8	1.2	0.6	0.8	0.8	0.8	1.4	1.2
Europe	(4.20)	(1.68)	(1.88)	(1.81)	(1.49)	(1.28)	(1.92)	(1.66)	(1.44)	(1.56)	(1.44)	(1.42)	(1.66)	(1.62)
France	3.2	1.4	1.6	1.6	1.2	1.2	1.6	1.6	1.4	1.4	1.4	1.6	1.2	1.4
Germany	4.0	1.6	2.0	1.6	1.2	1.4	2.0	1.6	1.4	1.2	1.4	1.6	1.6	1.8
Greece	5.2	1.2	1.8	2.2	1.8	1.6	2.4	2.2	2.0	2.0	1.8	1.8	2.0	2.0
Ireland	4.8	1.6	1.4	1.6	1.4	1.2	1.6	1.6	1.2	1.2	1.2	1.4	1.4	1.6
Italy	3.6	1.6	2.2	2.5	2.3	1.4	2.4	2.0	2.0	2.2	2.0	1.8	2.4	2.0
Netherlands	4.6	1.8	2.2	2.0	1.6	1.4	1.8	1.6	1.6	1.8	1.2	1.2	1.8	1.6
Portugal	4.2	2.0	2.0	2.2	2.0	1.8	2.2	1.6	1.4	1.8	2.0	1.8	1.8	1.8
Russia	4.0	1.8	2.2	1.6	1.2	0.8	1.0	0.8	0.6	1.0	1.2	1.0	1.2	0.6
Spain	4.0	2.0	1.8	1.6	1.2	1.2	2.4	2.0	1.8	1.6	1.2	1.2	1.8	2.0
UK	4.4	1.8	1.6	1.2	1.0	0.8	1.8	1.6	1.0	1.4	1.0	0.8	1.4	1.4
Asia	(4.89)	(2.00)	(1.74)	(1.36)	(1.01)	(0.60)	(1.03)	(0.86)	(0.66)	(0.86)	(0.97)	(0.66)	(1.06)	(0.86)
China	4.8	1.4	2.4	2.3	1.5	0.4	1.4	1.0	0.4	0.6	0.8	0.4	1.0	1.0
Hong Kong	4.0	1.8	2.0	1.8	1.8	1.4	1.4	1.4	1.4	1.4	1.6	1.6	1.6	1.4
India	4.4	3.0	1.6	1.0	0.8	0.2	1.2	0.8	0.4	0.6	1.0	0.4	1.0	0.6
Japan	4.6	2.6	1.4	1.4	1.0	0.8	1.4	1.2	1.0	1.0	1.2	1.0	1.6	1.2
South Korea	5.6	1.6	2.2	0.8	0.6	0.4	0.2	0.2	0.0	0.4	0.6	0.2	0.6	0.2
Taiwan	4.6	1.2	1.8	1.2	0.6	0.4	0.4	0.4	0.4	0.6	0.6	0.2	0.6	0.6
Thailand	6.2	2.4	0.8	1.0	0.8	0.6	1.2	1.0	1.0	1.4	1.0	0.8	1.0	1.0

Note that the values in (.) are the average VRate/α for each method/model in each region.

Table 5: VaR prediction performance using 2 HS methods and 12 risk models and 500 forecasted stock returns under $\alpha = 5\%$. VRate/α are given.

Markets	ST	LT	RM	G-n	G-t	G-sk	GJRn	GJRt	GJRsk	1STF GARCH	2STF GARCH	2STsk GARCH	ESTF GARCH	DT GARCH
Americas	(1.74)	(1.08)	(1.06)	(0.98)	(1.09)	(0.96)	(0.96)	(1.00)	(0.88)	(0.94)	(1.00)	(0.84)	(1.04)	(0.90)
Brazil	1.80	1.16	1.12	1.00	1.06	0.96	0.88	0.88	0.84	0.88	0.92	0.88	0.96	0.80
U.S.	1.68	1.00	1.00	0.96	1.12	0.96	1.04	1.12	0.92	1.00	1.08	0.80	1.12	1.00
Europe	(1.59)	(0.95)	(1.07)	(1.07)	(1.13)	(1.05)	(1.05)	(1.10)	(1.00)	(1.16)	(1.14)	(1.05)	(1.26)	(1.10)
France	1.60	0.88	1.00	1.12	1.20	1.00	1.08	1.08	1.00	1.20	1.24	1.16	1.20	1.20
Germany	1.68	0.92	0.92	1.08	1.08	1.00	1.04	1.04	1.00	1.20	1.16	1.08	1.20	1.24
Greece	1.68	1.12	1.16	1.08	1.16	1.28	1.24	1.28	1.28	1.36	1.12	1.24	1.64	1.12
Ireland	1.60	1.12	0.92	0.88	0.96	0.96	0.92	0.96	0.92	1.00	1.00	1.00	1.08	1.04
Italy	1.52	0.80	1.04	1.20	1.26	1.04	1.12	1.16	1.04	1.20	1.24	1.08	1.56	1.12
Netherlands	1.60	0.80	1.12	0.84	1.12	0.96	1.00	1.00	0.96	1.16	1.08	0.96	1.24	1.12
Portugal	1.64	1.08	1.12	1.24	1.16	1.24	1.16	1.28	1.16	1.24	1.24	1.08	1.24	1.08
Russia	1.40	1.00	1.16	0.96	1.00	0.80	0.76	0.84	0.76	0.80	1.00	0.80	0.84	0.84
Spain	1.72	1.00	1.04	1.12	1.12	1.12	1.12	1.24	1.04	1.12	1.12	0.96	1.28	1.12
UK	1.44	0.80	1.20	1.20	1.24	1.12	1.04	1.12	0.88	1.28	1.20	1.16	1.28	1.16
Asia	(1.66)	(1.10)	(1.06)	(0.95)	(1.11)	(0.88)	(0.83)	(0.95)	(0.82)	(0.93)	(1.01)	(0.87)	(0.97)	(0.97)
China	1.64	0.96	1.16	0.96	1.26	0.72	0.76	0.96	0.72	0.76	0.84	0.64	0.76	0.76
Hong Kong	1.56	1.04	1.04	0.96	1.04	0.92	0.88	0.92	0.88	0.96	1.12	0.96	0.96	0.84
India	1.48	1.16	1.08	1.08	1.24	0.92	0.84	0.84	0.76	0.84	1.04	0.80	0.84	0.96
Japan	1.52	1.24	0.92	0.92	0.92	0.88	1.04	1.04	0.92	0.92	0.96	0.88	1.08	1.12
South Korea	1.80	0.84	1.16	1.12	1.20	1.00	0.92	1.04	0.92	1.08	1.16	1.04	1.16	1.08
Taiwan	1.72	0.92	1.04	0.92	1.12	0.76	0.76	0.88	0.72	0.92	0.96	0.76	1.04	1.04
Thailand	1.92	1.52	1.04	0.68	1.00	0.96	0.64	1.00	0.84	1.00	0.96	1.00	0.96	0.96

Note that the values in (.) are the average VRate/α for each method/model in each region.

Table 6: Evaluating VaR prediction performance using the 19 markets with 500 out-of-sample forecasting

	Mean	Med	Ave loss*	Min	Max
1%					
ST	4.55	4.60	1.131	3.20	6.20
LT	1.78	1.60	1.008	1.20	3.00
RM	1.81	1.80	0.955	0.80	2.40
G-n	1.61	1.60	0.796	0.75	2.50
G-t	1.26	1.20	0.634	0.60	2.25
G-sk	0.98	1.00	0.475	0.20	1.80
GJRn	1.53	1.60	0.796	0.20	2.40
GJRt	1.26	1.40	0.635	0.20	2.20
GJRsk	1.04	1.00	0.476	0.00	2.00
1STF-GARCH	1.20	1.20	0.581	0.40	2.20
2STF-GARCH	1.20	1.20	0.581	0.60	2.00
2STsk-GARCH	1.06	1.00	0.475	0.20	1.80
ESTF-GARCH	1.36	1.40	0.688	0.40	2.40
DT-GARCH	1.25	1.40	0.635	0.20	2.00
5%					
ST	8.16	8.20	5.104	7.00	9.60
LT	5.09	5.00	2.111	4.00	7.60
RM	5.33	5.20	3.687	4.60	6.00
G-n	5.08	5.00	2.371	3.40	6.20
G-t	5.59	5.60	3.953	4.60	6.25
G-sk	4.89	4.80	1.318	3.60	6.40
GJRn	4.80	5.00	2.370	3.20	6.20
GJRt	5.18	5.20	2.635	4.20	6.40
GJRsk	4.62	4.60	1.054	3.60	6.40
1STF-GARCH	5.22	5.00	2.374	3.80	6.80
2STF-GARCH	5.38	5.40	3.162	4.20	6.20
2STsk-GARCH	4.81	4.80	1.844	3.20	6.20
ESTF-GARCH	5.64	5.60	3.436	3.80	8.20
DT-GARCH	5.16	5.40	3.161	3.80	6.20

*:

$$\Psi(\text{VRte}) = \begin{cases} \alpha + (\text{VRate} - \alpha)^2 & \text{if VRate} > \alpha \\ 0 & \text{if VRate} \leq \alpha \end{cases},$$

Table 7: Evaluating VaR estimates based on Lopez's loss functions using the 19 markets with 500 out-of-sample forecasting

	Mean	Med	Std	Min	Max(1)	Mean	Med	Std	Min	Max(2)
	Quadratic Loss					Absolute Loss				
1%										
ST	8.63	7.94	3.03	5.75	18.25	7.49	7.48	1.13	6.06	11.17
LT	3.05	2.57	1.73	1.55	9.25	3.24	2.56	2.08	1.78	11.04
RM	3.04	2.89	1.11	1.70	5.60	2.84	2.97	0.67	1.56	3.98
G-n	2.87	2.42	1.69	1.10	8.11	2.56	2.69	0.94	1.25	4.73
G-t	2.15	1.83	1.44	0.71	6.57	1.93	1.78	0.80	0.81	3.77
G-sk	1.73	1.39	1.41	0.33	6.21	1.53	1.60	0.82	0.36	3.44
GJRn	2.88	2.12	2.22	0.25	9.39	2.50	2.33	1.26	0.30	5.49
GJRt	3.38	1.84	6.17	0.21	28.13	2.11	2.03	1.43	0.24	6.66
GJRsk	1.90	1.36	1.87	0.00	7.25	1.61	1.52	1.07	0.00	4.25
1STF-GARCH	2.34	1.72	2.00	0.43	8.02	1.99	2.00	1.18	0.48	4.68
2STF-GARCH	2.21	1.52	1.64	0.68	7.20	1.94	1.67	0.90	0.79	3.93
2STsk-GARCH	1.83	1.36	1.55	0.23	6.75	1.63	1.56	0.94	0.27	3.71
ESTF-GARCH	2.65	2.00	2.30	0.46	9.64	2.25	1.98	1.20	0.53	5.08
DT-GARCH	2.32	1.72	1.92	0.42	7.67	1.98	1.92	1.04	0.49	4.47
	Mean	Med	Std	Min	Max(1)	Mean	Med	Std	Min	Max(3)
	Quadratic Loss					Absolute Loss				
5%										
ST	16.54	14.95	6.06	10.56	36.91	15.47	14.11	6.76	11.11	41.91
LT	9.97	8.78	4.04	5.64	20.07	8.65	8.32	1.79	6.06	12.44
RM	10.08	8.93	3.39	6.69	21.08	8.92	8.53	1.40	6.88	13.07
G-n	10.00	8.12	4.16	6.24	23.25	8.63	8.35	1.83	5.81	12.79
G-t	10.93	9.11	4.23	7.27	24.68	9.42	8.89	1.68	7.35	13.53
G-sk	9.30	7.77	4.31	4.95	24.18	8.15	7.80	1.94	5.49	13.97
GJRn	9.23	6.95	4.87	4.62	25.24	7.98	7.39	2.19	5.20	14.04
GJRt	10.00	8.78	4.95	5.15	26.31	8.65	8.04	2.13	5.79	14.57
GJRsk	8.86	7.74	4.86	4.40	25.60	7.65	6.96	2.18	4.90	14.27
1STF-GARCH	10.70	8.46	5.77	5.84	29.92	8.94	8.58	2.59	5.95	16.06
2STF-GARCH	10.79	9.03	4.40	6.76	24.57	9.25	8.78	1.81	7.31	13.33
2STsk-GARCH	9.28	7.46	4.42	4.92	24.28	8.06	8.10	2.01	5.47	13.74
ESTF-GARCH	11.69	9.09	6.49	6.43	32.97	9.73	9.20	3.06	6.24	18.29
DT-GARCH	10.37	8.53	4.86	5.81	26.04	8.74	8.04	2.00	6.02	13.50

(1): All extreme quadratic losses occurred in the Greece market, except for "LT" method.

(2): All extreme absolute losses occurred in the Greece market, except for "LT" and RiskMetrics.

(3): All extreme absolute losses occurred in the Greece market, except for the "ST" method.

Table 8: Counts of model rejections for three backtests across the 19 markets at the 5% level

	$\alpha = 1\%$				$\alpha = 5\%$			
Model	UC	CC	DQ_4	Total	UC	CC	DQ_4	Total
ST	19	19	19	19	19	19	19	19
LT	5	3	12	13	1	2	7	7
RM	9	2	14	15	0	2	4	4
G-n	5	0	7	9	0	0	0	0
G-t	3	1	4	6	1	1	0	1
G-sk	2	0	4	6	0	0	0	0
GJRn	4	2	3	5	0	0	0	0
GJRt	6	3	5	7	1	0	0	1
GJRsk	5	0	4	6	0	0	0	0
1STF-GARCH	2	0	3	4	0	0	0	0
2STF-GARCH	2	0	6	7	0	0	1	1
2STsk-GARCH	2	0	4	6	0	0	1	1
ESTF-GARCH	3	2	4	5	2	2	3	3
DT-GARCH	4	0	2	4	0	0	0	0

Table 9: Evaluating VaR prediction performance over the time period from January 2007 to December 2009 at the 1% and 5% levels.

Method/Model	Japan				U.S.			
	1% VRate/α	UC	CC	DQ	1% VRate/α	UC	CC	DQ
1%								
ST	4.17				5.33			
LT	2.50				3.17			
MR	2.18				2.80			
G-n	1.77	√	√		3.34			
G-t	1.64	√	√		1.87		√	
G-sk	1.50	√	√		1.60	√	√	
GJRn	1.91		√		3.20			
GJRt	1.64	√	√	√	1.87		√	
GJRsk	1.50	√	√	√	1.47	√	√	√
1STF-GARCH	2.18		√		2.80			
2STF-GARCH	1.50	√	√		2.67			
2STsk-GARCH	1.50	√	√		2.00			
ESTF-GARCH	3.68	√	√		3.74	√		
DT	1.77	√	√	√	2.54			
RV-n	2.86				2.94			
RV-t	2.73				2.40			
RV-sk	2.46				1.34	√	√	√
	5% VRate/α	UC	CC	DQ	5% VRate/α	UC	CC	DQ
5%								
ST	1.73				1.90		√	
LT	1.17	√			1.40		√	
MR	1.58				1.44			
G-n	1.64				1.55			
G-t	1.64				1.60			
G-sk	1.50				1.47			
GJRn	1.45				1.52			
GJRt	1.50			√	1.50			
GJRsk	1.39				1.42			
1STF-GARCH	1.58				1.44		√	
2STF-GARCH	1.66				1.60			
2STsk-GARCH	1.56				1.50			
ESTF-GARCH	2.02				2.00			
DT	1.53				1.60			
RV-n	2.18				1.63			
RV-t	2.21				1.60			
RV-sk	1.99				1.39			

"√" indicates that we fail to reject H_0 at the 5% significance level.

Table 10: Evaluating VaR estimates based on Lopez's loss functions over the time period from January 2007 to December 2009 at the 1% and 5% levels.

	Quadratic Loss		Absolute Loss	
	NK225	SP500	NK225	SP500
1%				
ST	14.19	12.41	8.41	9.79
LT	16.82	7.69	6.75	5.75
MR	8.53	4.85	4.94	4.42
G-n	6.54	5.60	3.99	5.13
G-t	5.94	3.44	3.68	3.08
G-sk	5.17	2.84	3.33	2.56
GJRn	4.84	5.04	3.65	4.70
GJRt	4.04	3.20	3.09	2.92
GJRsk	3.47	2.48	2.73	2.26
1STF-GARCH	4.31	4.56	3.49	4.14
2STF-GARCH	5.75	4.21	3.57	3.98
2STsk-GARCH	5.14	3.16	3.33	3.03
ESTF-GARCH	5.88	6.71	3.17	6.30
DT	6.35	4.28	5.45	4.05
RV-n	7.34	5.94	5.68	4.72
RV-t	8.47	4.55	7.26	3.73
RV-sk	7.39	2.81	6.54	2.26
	Quadratic Loss		Absolute Loss	
	NK225	SP500	NK225	SP500
5%				
ST	28.93	23.82	16.82	18.14
LT	55.01	24.82	16.73	14.06
MR	26.32	17.11	15.13	13.61
G-n	24.12	19.25	15.66	14.89
G-t	25.72	19.77	15.98	15.34
G-sk	23.39	17.60	14.52	14.01
GJRn	20.15	17.17	13.86	13.90
GJRt	21.19	17.00	14.31	13.82
GJRsk	19.24	14.90	13.05	12.60
1STF-GARCH	22.27	18.63	15.03	13.81
2STF-GARCH	26.43	19.69	16.47	15.50
2STsk-GARCH	24.20	17.53	15.14	14.28
ESTF-GARCH	24.82	25.71	15.40	19.51
DT	24.49	19.30	17.12	15.13
RV-n	31.80	18.46	20.88	14.29
RV-t	34.45	18.08	23.46	14.05
RV-sk	30.67	15.43	21.27	11.99

Table 11: Evaluation of volatility forecasting based on three proxies and three loss functions

	RMSE			MAD			QLIKE		
Proxy 1	Mean	Med	Std	Mean	Med	Std	Mean	Med	Std
RM	0.895	0.859	0.257	0.726	0.709	0.212	0.490	0.499	0.096
G-n	0.891	0.849	0.243	0.732	0.680	0.200	0.498	0.503	0.102
G-t	0.890	0.847	0.243	0.731	0.680	0.200	0.496	0.504	0.102
G-sk	1.279	1.125	0.723	0.998	0.799	0.581	0.653	0.573	0.307
GJRn	1.295	1.146	0.707	0.991	0.856	0.565	0.603	0.571	0.103
GJRt	1.313	1.162	0.716	1.001	0.849	0.573	0.620	0.573	0.174
GJRsk	1.302	1.152	0.711	0.992	0.844	0.569	0.626	0.556	0.233
1STF-GARCH	0.890	0.848	0.241	0.728	0.682	0.197	0.497	0.510	0.103
2STF-GARCH	0.889	0.859	0.243	0.729	0.707	0.199	0.494	0.506	0.102
2STsk-GARCH	1.301	1.111	0.700	1.013	0.787	0.563	0.604	0.591	0.094
ESTF-GARCH	0.878	0.838	0.230	0.716	0.684	0.185	0.491	0.493	0.101
DT-GARCH	0.882	0.834	0.245	0.720	0.684	0.202	0.489	0.498	0.103
Proxy 2	Mean	Med	Std	Mean	Med	Std	Mean	Med	Std
RM	0.647	0.583	0.241	0.543	0.486	0.233	0.366	0.132	0.851
G-n	0.645	0.583	0.231	0.558	0.509	0.237	0.379	0.140	0.866
G-t	0.643	0.589	0.232	0.556	0.515	0.237	0.378	0.141	0.867
G-sk	1.151	0.917	0.752	0.900	0.692	0.637	0.498	0.218	0.959
GJRn	1.184	0.890	0.727	0.900	0.690	0.612	0.452	0.221	0.933
GJRt	1.206	0.926	0.734	0.911	0.704	0.618	0.501	0.226	0.948
GJRsk	1.193	0.923	0.731	0.902	0.695	0.615	0.511	0.220	0.975
1STF-GARCH	0.649	0.607	0.245	0.558	0.509	0.251	0.380	0.142	0.877
2STF-GARCH	0.646	0.595	0.228	0.555	0.497	0.235	0.377	0.138	0.867
2STsk-GARCH	1.178	0.920	0.725	0.916	0.690	0.616	0.452	0.225	0.929
ESTF-GARCH	0.630	0.611	0.231	0.541	0.493	0.237	0.374	0.139	0.867
DT-GARCH	0.637	0.600	0.246	0.549	0.496	0.248	0.373	0.143	0.872
Proxy 3	Mean	Med	Std	Mean	Med	Std	Mean	Med	Std
RM	0.611	0.517	0.246	0.498	0.398	0.240	0.310	0.106	0.829
G-n	0.606	0.541	0.242	0.510	0.447	0.246	0.320	0.109	0.844
G-t	0.604	0.540	0.243	0.508	0.444	0.247	0.319	0.109	0.845
G-sk	1.097	0.853	0.723	0.844	0.647	0.611	0.465	0.179	0.947
GJRn	1.129	0.829	0.699	0.848	0.635	0.589	0.416	0.187	0.912
GJRt	1.150	0.861	0.706	0.859	0.649	0.595	0.472	0.181	0.932
GJRsk	1.138	0.859	0.704	0.849	0.640	0.592	0.483	0.176	0.965
1STF-GARCH	0.613	0.551	0.258	0.511	0.467	0.262	0.322	0.113	0.855
2STF-GARCH	0.607	0.537	0.239	0.508	0.447	0.244	0.319	0.113	0.845
2STsk-GARCH	1.120	0.860	0.697	0.858	0.627	0.592	0.410	0.184	0.909
ESTF-GARCH	0.600	0.508	0.246	0.500	0.443	0.248	0.319	0.119	0.845
DT-GARCH	0.597	0.523	0.257	0.500	0.450	0.260	0.410	0.184	0.909