Bayesian Time-Varying Quantile Forecasting for Value-at-Risk in Financial Markets

Richard H. Gerlach
Discipline of Operations Management and Econometrics, University of Sydney, NSW 2006, Australia (*richard.gerlach@sydney.edu.au*)

Cathy W. S. Chen
Department of Statistics, Feng Chia University, 407, Taiwan (*chenws@fcu.edu.tw*)

Nancy Y. C. Chan
Department of Statistics, Feng Chia University, 407, Taiwan (*ycchan@fcu.edu.tw*)

Recently, advances in time-varying quantile modeling have proven effective in financial Value-at-Risk forecasting. Some well-known dynamic conditional autoregressive quantile models are generalized to a fully nonlinear family. The Bayesian solution to the general quantile regression problem, via the Skewed-Laplace distribution, is adapted and designed for parameter estimation in this model family via an adaptive Markov chain Monte Carlo sampling scheme. A simulation study illustrates favorable precision in estimation, compared to the standard numerical optimization method. The proposed model family is clearly favored in an empirical study of 10 major stock markets. The results that show the proposed model is more accurate at Value-at-Risk forecasting over a two-year period, when compared to a range of existing alternative models and methods.

KEY WORDS: Asymmetric; CAViaR model; GARCH; Regression quantile; Skewed-Laplace distribution.

1. INTRODUCTION

Value-at-Risk (VaR) is the benchmark for risk measurement and subsequent capital allocation for financial institutions worldwide, as chosen by the Basel Committee on Banking Supervision (see Basel II: *http://www.bis.org/publ/bcbsca.htm*). VaR is a one-to-one function of a quantile, over a given time interval, of an asset portfolio's conditional return distribution (see Jorion 1996). Many competing models/methods have been used in the literature to estimate and forecast quantiles, and hence VaR (see Kuester, Mittnik, and Paolella 2006 for a review). We focus on the conditional autoregressive VaR (CAViaR) models proposed by Engle and Manganelli (2004), to estimate quantiles (VaR) directly, and expand the existing CAViaR models into a fully nonlinear family of dynamic models, in the spirit of threshold GARCH modeling (see Zakoian 1994 and Brooks 2001). Such models capture a range of nonlinear and asymmetric behavior, including the well-known leverage effect, or volatility asymmetry, discovered by Black (1976).

Manganelli and Engle (2004) find that the classical numerical optimization procedure for quantile models of Koenker and Bassett (1978) is inefficient and problematic at very low quantiles, and presumably at low sample sizes as well, for CAViaR models. This article develops a Bayesian estimator for the proposed nonlinear CAViaR model family, adapted from the MCMC sampling scheme of Chen and So (2006). This MCMC estimator takes advantage of the gains in efficiency that numerical integration, for estimating the posterior mean, can display over standard numerical differentiation based optimisation, for example, in Engle and Manganelli (2004), especially in models with many parameters. The MCMC scheme is adaptive, expanding the MCMC methods for general quantile regression models in Yu and Moyeed (2001), Tsionas (2003), and Geraci and Bottai (2007).

Parametric GARCH-type models (see, e.g., Engle 1982 and Bollerslev 1986), with specified error distributions, are also well suited to quantile forecasting and are often used as benchmarks in VaR studies (e.g., Kuester, Mittnik, and Paolella 2006). However, some financial institutions employ sample return quantiles, called historical simulation (HS), for nonparametric estimation of VaR. Manganelli and Engle (2004) find that the basic CAViaR model outperforms standard GARCH and HS approaches for forecasting VaR from simulated data, especially for fat-tailed error processes. We show that this result also holds for some real market return data, at the 1% quantile level. Giacomini and Komunjer (2005) develop an encompassing test for forecast models and find that a CAViaR model is most useful at the 1% quantile level, but that a simple GARCH model with Gaussian errors outperforms the CAViaR model at the 5% level, for the U.S. S&P500 return data from 1985–2001. We extend these results in this article across 10 international market indices, finding that this conclusion generalizes across these markets, in a more recent time period. Further, Geweke and Keane (2007) propose a smoothly mixing regression approach to semiparametrically and flexibly estimate predictive densities, a method that can capture some types of heteroscedasticity. This approach outperforms GARCH models for the U.S. S&P500 index from 1995–1999. We consider all the above-mentioned VaR forecasting approaches in this article.

Journal of Business & Economic Statistics
October 2011, Vol. 29, No. 4
DOI: 10.1198/jbes.2010.08203

The proposed full model and MCMC methods are examined first through a simulation study, and second through application to various financial market stock indices in a study of VaR forecasting. The simulation study illustrates favorable estimation performance, in terms of precision and efficiency, compared with numerically optimizing the usual quantile criterion function. The empirical study illustrates that CAViaR models perform favorably compared to RiskMetrics and general GARCH estimators, for VaR forecasting, during the period 2005–2007, especially at the 1% quantile level; while simpler models, including GARCH and RiskMetrics do comparably well at the 5% quantile level.

The article is organized as follows: Section 2 discusses dynamic quantile estimation; Section 3 introduces the proposed family of dynamic autoregressive quantile models; Section 4 presents the MCMC methods employed; Sections 5 and 6, respectively, present the simulation and empirical studies; and some concluding remarks are given in Section 7.

2. DYNAMIC QUANTILE AND VALUE-AT-RISK

This section discusses the general dynamic quantile problem and the relation to the Skewed Laplace (SL) distribution. The general dynamic quantile model may be written

$$y_t = f_t(\boldsymbol{\beta}, \mathbf{x}_{t-1}) + u_t, \tag{1}$$

where y_t is the observation at time t; $\mathbf{x}_{t-1}$ are the explanatory variables; $\boldsymbol{\beta}$ are unknown parameters and u_t is an error term. The function $f_t(\cdot)$ defines the dynamic link between y_t and $\mathbf{x}_{t-1}$ and is usually linear in $\boldsymbol{\beta}$ and $\mathbf{x}$, an aspect that is extended here. The conditional α level quantile is then

$$q_\alpha(y_t|\boldsymbol{\beta}, \mathbf{x}_{t-1}) = f_t(\boldsymbol{\beta}_\alpha, \mathbf{x}_{t-1}),$$

where $\boldsymbol{\beta}_\alpha$ is the solution to

$$\min_{\boldsymbol{\beta}} \sum_t \rho_\alpha(y_t - f_t(\boldsymbol{\beta}, \mathbf{x}_{t-1})). \tag{2}$$

The function $\rho(\cdot)$ is a loss function, specified as $\rho_\alpha(u) = u(\alpha - I(u < 0))$.

2.1 The Skewed-Laplace Connection

Yu and Moyeed (2001) and Tsionas (2003) illustrate the link between the solution to the quantile estimation problem and the likelihood for the SL distribution, as follows. The SL location-scale family, denoted $\mathrm{SL}(\mu, \tau, \alpha)$, has density function

$$p_\alpha(u) = \frac{\alpha(1-\alpha)}{\tau} \exp\left[-\rho_\alpha\left(\frac{u-\mu}{\tau}\right)\right], \tag{3}$$

where μ is the mode and $\tau > 0$ is a scale parameter. If it is assumed, in Model (1), that $u_t \sim \mathrm{SL}(0, \tau, \alpha)$ and is iid, then the likelihood function becomes

$$L_\alpha(\boldsymbol{\beta}, \tau; \mathbf{y}, \mathbf{X}) \propto \tau^{-n} \exp\left\{-\tau^{-1} \sum_{t=1}^{n} (y_t - f_t(\boldsymbol{\beta})) \times \left[\alpha - I_{(-\infty,0)}(y_t - f_t(\boldsymbol{\beta}))\right]\right\}. \tag{4}$$

Since (2) is contained in the exponent of the likelihood, the maximum likelihood estimate for $\boldsymbol{\beta}$ is equivalent to the quantile estimator in (2).

It is important to emphasize that, though we treat (4) as a likelihood function, the assumption that $\mathbf{y}$ follows an $\mathrm{SL}(\mu, \tau, \alpha)$ distribution is not used to parametrically estimate VaR. In practice, the parameter α is fixed and known during parameter estimation and only that single quantile of the distribution of y_t is estimated. Equation (4) is only employed as it leads to (maximum likelihood) estimation that is mathematically equivalent to (2). Use of (4) as a likelihood function, then allows a Bayesian approach to consider powerful computational estimation methods, such as adaptive MCMC algorithms, that employ numerical integration (which can be made arbitrarily accurate), instead of numerical optimization. Authors such as Yu and Moyeed (2001) and Tsionas (2003) have illustrated that accurate Bayesian estimation and inference is achieved under this approach. We further investigate and add to these findings for dynamic nonlinear quantile models.

2.2 Value-at-Risk

The Basel Capital Accord, originally signed by the Group of Ten countries in 1988, requires Authorized Deposit-taking Institutions to hold sufficient capital to provide a cushion against unexpected losses. VaR forecasts the minimum expected loss over a given time interval, at a given confidence level α (Jorion 1996). That is,

$$\alpha = \Pr(y_t < -\mathrm{VaR}_t|\mathbf{y}^{1,t-1}).$$

VaR is thus proportional to a quantile in the conditional one-step-ahead forecast distribution.

As in Manganelli and Engle (2004) VaR estimation methods can be classified as parametric, with a full distributional and model form specified; nonparametric, with minimal or no distributional or model assumptions; and semiparametric, where some assumptions are made, either about the error distribution, or the model dynamics, but not both (e.g., quantile regression and CAViaR models). Monte Carlo (MC) simulation methods are commonly used in all three categories.

A popular parametric method for VaR estimation is RiskMetrics (RMetrics), proposed by J. P. Morgan in 1996, where an IGARCH(1, 1) process, with no mean equation, is employed. For expositional purposes, the standard GARCH(1, 1) model is

$$y_t = \mu + a_t; \qquad a_t = \sigma_t \epsilon_t, \qquad \epsilon_t \overset{\text{iid}}{\sim} D(0, 1),$$
$$\sigma_t^2 = \alpha_1 + \alpha_2 \sigma_{t-1}^2 + \alpha_3 a_{t-1}^2,$$

with the RMetrics model a special case that sets $D \equiv \mathrm{N}(0, 1)$, $\mu = \alpha_0 = 0$, $\alpha_1 = 1 - \beta_1$, and $\beta_1 = 0.94$ (for daily data); and is thus nonstationary in volatility. Standard GARCH theory (e.g., see Tsay 2005) allows closed form solutions to the one-step-ahead quantiles of $y_t|\mathbf{y}^{1,t-1}$, given the parameter values and based on parametric errors. From here on, the standard GARCH(1, 1) with Gaussian errors is labeled GARCH-n, while the GARCH(1, 1) with standardized Student-t errors is denoted GARCH-t.

2.3 CAViaR Models for VaR

Engle and Manganelli (2004) propose various dynamic quantile functions $f(\cdot)$, which they called CAViaR models. We initially discuss three of their specifications:

Indirect GARCH(1, 1) (IG):

$$f_t(\boldsymbol{\beta}) = [\beta_1 + \beta_2 f_{t-1}^2(\boldsymbol{\beta}) + \beta_3 y_{t-1}^2]^{1/2}. \tag{5}$$

This equation is exactly equivalent to the dynamic quantile function for a GARCH(1, 1) model with an iid symmetric error distribution and mean $\mu = 0$. The model thus allows efficient estimation for GARCH(1, 1) quantiles with unspecified error distribution. This is an advantage: GARCH models are typically estimated under a parametric error distribution. However, it is well known that standard GARCH models tend to overreact to large return shocks (since they are squared). As such we prefer the absolute value model types below:

Symmetric Absolute Value (SAV):

$$f_t(\boldsymbol{\beta}) = \beta_1 + \beta_2 f_{t-1}(\boldsymbol{\beta}) + \beta_3 |y_{t-1}|. \tag{6}$$

The quantile again responds symmetrically to the lagged return y_{t-1}. This equation is equivalent to the quantile function for a standard deviation GARCH model (see, e.g., Zakoian 1994), where

$$\sigma_t = \gamma_1 + \gamma_2 \sigma_{t-1} + \gamma_2 |a_{t-1}|.$$

The two CAViaR models SAV and IG have symmetric responses to positive and negative observations. To account for financial market asymmetry, via the leverage effect (Black 1976), the SAV model is extended in Engle and Manganelli (2004) to

Asymmetric Slope (AS):

$$f_t(\boldsymbol{\beta}) = \beta_1 + \beta_2 f_{t-1}(\boldsymbol{\beta}) + \left(\beta_3 I_{(y_{t-1}>0)} + \beta_4 I_{(y_{t-1}<0)}\right)|y_{t-1}|, \tag{7}$$

where the dynamic quantile function can respond differently to positive and negative responses. Such a threshold nonlinear model is similar in spirit to the GJR-GARCH model (Glosten, Jagannathan, and Runkle 1993) or EGARCH (Nelson 1991), where asymmetry is modeled by adding one parameter only, and hence the types of asymmetry captured are limited. Again the AS model corresponds to a standard deviation GJR-GARCH model with mean $\mu = 0$ where

$$\sigma_t = \delta_1 + \delta_2 \sigma_{t-1} + \delta_3 I_{(a_{t-1}>0)} |a_{t-1}| + \delta_4 I_{(a_{t-1}<0)} |a_{t-1}|.$$

We expand these CAViaR models in Section 3 to capture more flexible asymmetric and nonlinear responses, via more general threshold nonlinear forms.

3. PROPOSED NONLINEAR DYNAMIC QUANTILE FAMILY

A popular model for explaining asymmetry in the mean is the threshold autoregressive (TAR) model of Tong (1978, 1983). As a natural extension of this model, Li and Li (1996) and Brooks (2001) modeled mean and volatility asymmetry together, expanding the first generation ARCH and GARCH models to be fully threshold nonlinear. In the same spirit, it is natural to extend the SAV and AS (CAViaR) models to the following threshold CAViaR model:

Threshold CAViaR (T-CAViaR):

$$f_t(\boldsymbol{\beta}) = \begin{cases} \beta_1 + \beta_2 f_{t-1}(\boldsymbol{\beta}) + \beta_3 |y_{t-1}|, & z_{t-1} \le r \\ \beta_4 + \beta_5 f_{t-1}(\boldsymbol{\beta}) + \beta_6 |y_{t-1}|, & z_{t-1} > r. \end{cases} \tag{8}$$

Here $\mathbf{z}$ is an observed threshold variable which could be exogenous, or self-exciting, that is, $z_t = y_t$ and r is the threshold value, typically set as $r = 0$, or estimated, though empirically many estimates in the literature are not significant from zero; as such we fix $r = 0$, which also makes the T-CAViaR a direct extension of the AS-CAViaR model in (7). We consider both exogenous and self-exciting threshold variables in this article. The exogenous threshold is the return on the U.S. S&P500 index and the model using that threshold is denoted as TCAVx. The self-exciting model is denoted TCAV.

Here each parameter in the dynamic quantile function can respond differently to positive and negative responses. We call this the T-CAViaR family since it includes the SAV ($r = \infty$) and AS ($r = 0$, $\beta_4 = \beta_1$ and $\beta_5 = \beta_2$) CAViaR models as special cases. Once again, this model is the semiparametric equivalent of the standard deviation T-GARCH model with mean $\mu = 0$

$$\sigma_t = \begin{cases} \delta_1 + \delta_2 \sigma_{t-1} + \delta_3 |a_{t-1}|, & z_{t-1} \le 0 \\ \delta_4 + \delta_5 \sigma_{t-1} + \delta_6 |a_{t-1}|, & z_{t-1} > 0. \end{cases} \tag{9}$$

We choose to focus on this absolute value CAViaR model-type. However, a corresponding T-CAViaR-IG model could be specified as

Threshold Indirect-GARCH (T-CAViaR-IG):

$$f_t(\boldsymbol{\beta}) = \begin{cases} [\beta_1 + \beta_2 f_{t-1}^2(\boldsymbol{\beta}) + \beta_3 y_{t-1}^2]^{1/2}, & z_{t-1} \le r \\ [\beta_4 + \beta_5 f_{t-1}^2(\boldsymbol{\beta}) + \beta_6 y_{t-1}^2]^{1/2}, & z_{t-1} > r. \end{cases} \tag{10}$$

We limit focus in this article to the T-CAViaR model family (8). These models are nonparametric in their error specifications, but simply extend the existing forms for the dynamic function $f(\cdot)$ to be fully threshold nonlinear: that is, they are semiparametric. Yu, Li, and Jin (2010) extend CAViaR using two approaches, namely the threshold and mixture type indirect-VaR models.

Dynamic models typically have constraints or restrictions on the parameters for stationarity (or positivity of dynamic variances). However, such restrictions are difficult to locate or derive for CAViaR models, and we choose not to set any in this article, as in Engle and Manganelli (2004).

4. BAYESIAN METHODS

Bayesian methods generally require the specification of a likelihood function and a prior distribution. The likelihood function for the T-CAViaR is completely specified by (4) and (8). We now specify the prior distribution.

4.1 Prior and Posterior Densities

We choose the prior to be uninformative over the possible region for the regression-type parameters $\boldsymbol{\beta}$. The joint prior is thus

$$\pi(\boldsymbol{\beta}) \propto 1,$$

which is equivalent to a flat prior on $\boldsymbol{\beta}$ over the real line, in six dimensions. Yu and Moyeed (2001) showed that the posterior in $\boldsymbol{\beta}$ is proper, under this improper prior, for general quantile regression models.

Using (4), (8), plus the standard Jeffreys prior $\pi(\tau) \propto \tau^{-1}$, the joint posterior density for $\boldsymbol{\beta}, \tau|\mathbf{y}$ is

$$p(\boldsymbol{\beta}, \tau|\mathbf{y}) \propto p(\boldsymbol{\beta}, \tau) L_\alpha(\boldsymbol{\beta}, \tau; \mathbf{y}) \\ \propto \tau^{-(n+1)} \exp\left\{-\tau^{-1} \sum_{t=2}^{n} \rho_\alpha(y_t - f_t(\boldsymbol{\beta}))\right\}, \quad (11)$$

which is in the form of an inverse gamma density in τ. Since estimation of τ is not relevant in VaR forecasting, this parameter is integrated out of (11), to obtain the marginal posterior for $\boldsymbol{\beta}|\mathbf{y}$:

$$p(\boldsymbol{\beta}|\mathbf{y}) = \int p(\boldsymbol{\beta}, \tau|\mathbf{y})\, d\tau \\ \propto \left[\sum_{t=2}^{n} \rho_\alpha(y_t - f_t(\boldsymbol{\beta}))\right]^{-n}, \quad (12)$$

using the fact that the inverse gamma density integrates to 1. This posterior is not in a form permitting direct inference on $\boldsymbol{\beta}$. We thus turn to computational methods for estimation and inference.

4.2 Adaptive MCMC Sampling Using Metropolis Methods

Sampling from $p(\boldsymbol{\beta}|\mathbf{y})$ directly is not possible given the nonstandard form, so instead a dependent (Markov chain) Monte Carlo sample is obtained from (12) via the Metropolis and Metropolis–Hastings (MH) (Metropolis et al. 1953; Hastings 1970) algorithms. To speed convergence and to allow desirable mixing properties in the chain, an adaptive MCMC algorithm for $\boldsymbol{\beta}|\mathbf{y}$ is employed combining a random walk Metropolis (RW-M) and an independent kernel (IK-) MH algorithm.

We modify the adaptive sampling scheme of Chen and So (2006), who used Gaussian proposal densities. Such a proposal can get "stuck" in local modes and take a large number of iterates to move out of that area of the posterior. To improve on this aspect, for the burn-in period iterations, a Student-t proposal distribution, with low degrees of freedom (e.g., df $= 5$), is employed in a RW-M algorithm. The scale matrix, which can be chosen as diagonal with positive values, is subsequently tuned to achieve optimal acceptance rates of between 20% and 50%, based on recommendations in Gelman, Roberts, and Gilks (1996) and Chen and So (2006). After the burn-in period, the sample mean vector and sample variance–covariance matrix are formed using these M iterates of $\boldsymbol{\beta}$. These are subsequently employed in the sampling period (iterations $M + 1$ to N) as the mean and scale matrix for another Student-t proposal distribution in an IK-MH algorithm.

This adaptive proposal updating procedure should speed mixing in the posterior distribution, over that for the simple RW-M method, as long as the burn-in period has "covered" the posterior distribution. The Student-t proposals employed here will further assist in achieving coverage and mixing, over the Gaussian, for both the burn-in and sampling periods, by lowering the probability of getting stuck in local modes for long periods and allowing for occasional large jumps around the posterior space. We extensively examined trace plots and autocorrelation function (ACF) plots, from multiple runs of the MCMC sampler, from differing starting points, for each element of $\boldsymbol{\beta}$, so as to confirm convergence and to infer adequate coverage. We also employed Gelman's 'R' statistic (see Gelman et al. 2005, p. 296), which assesses speed of mixing and efficiency of convergence. The values obtained for real and simulated data are typically very close to 1, and almost always below 1.05, implying fast mixing and good convergence properties, for all parameters. Observed MH acceptance rates for this scheme are usually from 15%–40%.

5. SIMULATION STUDY

A simulation study is performed to study the comparative performance of the proposed MCMC method and the classical quantile estimator (2), in terms of parameter and quantile estimation and forecasting. The aim is to partially illustrate similar lack of bias for these two methods, while highlighting the increased precision of the MCMC estimator. The results presented focus on the full T-CAViaR model only; results employing the AS and SAV models are available from the authors on request: they are very similar to those presented here.

A specific choice of parametric error distribution is needed here. We choose the equivalent, to the full T-CAViaR model, standard deviation T-GARCH in (9). Samples of size $n = 2000$ were simulated from this model, with Student-t errors, specified as

$$y_t = a_t,$$
$$a_t = \sigma_t \varepsilon_t, \qquad \varepsilon_t \overset{\text{iid}}{\sim} t_6^*,$$
$$\sigma_t = \begin{cases} 0.2 + 0.03|a_{t-1}| + 0.95\sigma_{t-1}, & \text{if } a_{t-1} \le 0 \\ 0.05 + 0.15|a_{t-1}| + 0.75\sigma_{t-1}, & \text{if } a_{t-1} > 0. \end{cases}$$

Here t_6^* represents the standardized Student-t distribution with $\nu = 6$. The true one-step-ahead α quantile is then $q_\alpha(y_{t+1}|\boldsymbol{\beta}) = \sigma_{t+1} T_6^{-1}(\alpha)\sqrt{\frac{4}{6}}$, where T_6^{-1} is the inverse Student-t cdf.

400 datasets were simulated. The full T-CAViaR model is fit to each, once using the MCMC method in Section 4 and secondly using the classical estimator, employing the 'fminsearch' routine in Matlab software, to numerically minimize (2). The Matlab code employed is adapted, and updated for the full T-CAViaR model, from freely available code written by Simone Manganelli; downloadable from *http://www.simonemanganelli.org/Simone/Research.html*. The MCMC sample size is $N = 40{,}000$, with a burn-in of $M = 15{,}000$, iterations. Initial values were randomly set inside (0, 1) for each parameter. For the classical estimator, a grid of starting values were chosen for each parameter (from $[-1, 1]$), then parameter estimates, starting from each set of grid values, were found; the global minimum value of (2) gave the final classical estimates. VaR estimates in-sample, as well as a one-step-ahead forecast for VaR at time $n + 1 = 2001$ were also calculated. Following Basel II risk management guidelines, quantile levels of $\alpha = 0.01, 0.05$ were considered.

Estimation results are summarized in Table 1. The corresponding parameters of the T-CAViaR model are: $\beta_1(\alpha) =$

Table 1. Summary statistics for the two estimators of the T-CAViaR model using data simulated from a T-GARCH-t model

	$\alpha = 1\%$			$\alpha = 5\%$		
	True	Mean	Std.	True	Mean	Std.
MCMC par.						
β_1	−0.513	−0.532	**0.524**	−0.317	−0.333	0.180
β_2	0.95	**0.946**	**0.147**	0.95	0.941	0.084
β_3	−0.077	−0.079	**0.224**	−0.048	**−0.047**	**0.073**
β_4	−0.128	**−0.231**	**0.460**	−0.079	−0.106	**0.148**
β_5	0.75	**0.719**	**0.145**	0.75	−0.734	**0.080**
β_6	−0.385	−0.407	**0.190**	−0.238	−0.255	**0.079**
$\beta_1 - \beta_4$	−0.385	**−0.301**	**0.853**	−0.238	**−0.227**	**0.290**
$\beta_2 - \beta_5$	0.20	**0.227**	**0.233**	0.20	**0.207**	**0.133**
$\beta_3 - \beta_6$	0.308	0.329	**0.279**	0.190	0.208	**0.106**
$\widehat{\mathrm{VaR}}_{n+1}$	−3.909	**−3.922**	1.449	−2.417	**−2.415**	**0.845**
VaR_{n+1}		−3.909	1.349		−2.417	0.834
$\widehat{\mathrm{VaR}}_{n+1} - \mathrm{VaR}_{n+1}$	0	**−0.013**	0.631	0	**0.002**	**0.223**
MAD		**0.432**	**0.154**		**0.162**	**0.054**
MedAD		**0.327**	**0.111**		**0.124**	**0.045**
RMSE		**0.572**	**0.285**		**0.207**	**0.072**
Quantile par.						
β_1	−0.513	**−0.496**	0.541	−0.317	**−0.320**	**0.179**
β_2	0.95	0.955	0.153	0.95	**0.948**	**0.083**
β_3	−0.077	**−0.076**	0.233	−0.048	−0.041	0.074
β_4	−0.128	−0.246	0.529	−0.079	**−0.096**	0.152
β_5	0.75	0.718	0.158	0.75	**0.743**	0.081
β_6	−0.385	**−0.402**	0.195	−0.238	**−0.248**	**0.079**
$\beta_1 - \beta_4$	−0.385	−0.250	0.938	−0.238	−0.225	0.303
$\beta_2 - \beta_5$	0.20	0.237	0.252	0.20	**0.205**	0.139
$\beta_3 - \beta_6$	0.308	**0.327**	0.288	0.190	0.207	0.108
$\widehat{\mathrm{VaR}}_{t+1}$	−3.909	−3.924	**1.443**	−2.417	−2.401	0.854
VaR_{t+1}		−3.909	1.349		−2.417	0.834
$\widehat{\mathrm{VaR}}_{n+1} - \mathrm{VaR}_{n+1}$	0	−0.015	**0.629**	0	0.017	0.233
MAD		0.445	0.162		0.166	0.056
MedAD		0.336	0.117		0.127	0.046
RMSE		0.590	0.293		0.213	0.076

$0.2T_6^{-1}(\alpha)\sqrt{\frac{4}{6}}$; $\beta_2 = 0.95$; $\beta_3(\alpha) = 0.03T_6^{-1}(\alpha)\sqrt{\frac{4}{6}}$; while $\beta_4(\alpha) = 0.05T_6^{-1}(\alpha)\sqrt{\frac{4}{6}}$; $\beta_5 = 0.75$, and $\beta_6(\alpha) = 0.15T_6^{-1} \times (\alpha)\sqrt{\frac{4}{6}}$; giving the true parameter values in Table 1.

Table 1 reports the average of the 400 estimates for each parameter and their standard deviation (Std.), from each true value, for both methods. These summaries are also shown for the true (simulated) VaR at $t = n + 1 = 2001$, the MCMC one-step-ahead forecast VaR for $t = 2001$, and the difference between forecast and true VaR by MCMC. Finally, the measures mean absolute deviation (MAD), median absolute deviation (MedAD) and root mean square error (RMSE) were used to assess the competing in-sample estimates of VaR.

We consider estimation bias and precision. Regarding bias, both sets of parameter estimates average close to the true values across both quantile levels. Values highlighted in bold are closer to the true value. Clearly, the methods are comparable regarding minimal estimation bias. However, regarding precision, the MCMC estimates are almost all closer in squared error terms to the true value than the standard quantile estimators. While the differences are mostly marginal, the agreement across parameters in this aspect is striking. Clearly the MCMC estimator is consistently more precise than the classical estimator for this model, over the 400 datasets generated, at $n = 2000$. We have found this result also holds for different parameter value sets and across the AS and SAV models also. Further, this result also carries over to VaR estimation in sample, with all accuracy measures favouring the MCMC estimates. However, one-step-ahead forecast VaR results are mixed and comparable between the two methods. In short, the MCMC method should be the preferred estimator, as it gives comparable and minimal levels of bias and marginally but clearly better precision, thus better efficiency, in parameter and in-sample VaR estimation at this sample size.

This is an interesting result given that asymptotically, Bayesian estimation under flat priors and classical estimates should yield the same results. Indeed, for $\alpha = 0.01$, the precision of the MCMC estimates and MLEs do converge by about $n = 10{,}000$ for this model, while for $\alpha = 0.05$ convergence is clear by around $n = 5000$, as found by further simulations not reported here. Instead of speculating on the reasons for this observed difference in precision, we leave a more thorough study of this question for future research.

6. TESTING VAR MODELS

A common nontest criterion to compare VaR models is the rate of violation, defined as the proportion of observations for which the actual return is more extreme than the forecasted VaR level, over the forecast period. The violation rate is

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

where n is the learning sample size and m is the forecast or test sample size. A forecast model's VRate should be close to the nominal level α. We employed the ratio VRate/α, to help compare the competing models, where models with $\mathrm{VRate}/\alpha \approx 1$ are most desirable. When $\mathrm{VRate} < \alpha$, risk and loss estimates are conservative (higher than actual), while alternatively, when $\mathrm{VRate} > \alpha$, risk estimates are lower than actual and financial institutions may not allocate sufficient capital to cover likely future losses. Here solvency outweighs profitability and for models where VRate/α are equidistant from 1, lower or conservative rates are preferred; for example, $\mathrm{VRate}/\alpha = 0.9$ is preferred to $\mathrm{VRate}/\alpha = 1.1$, as in Wong and So (2003).

We further consider three standard hypothesis-testing 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). Both the CC and DQ are joint tests where the null hypothesis consists of: independence of a model's violations, equivalently correct conditional violation rate for a given model, combined with a correct UC rate. The DQ test is well known to be more powerful than the CC test; see Berkowitz, Christoffersen, and Pelletier (2010). These tests are now standard; we refer readers to the original papers for details.

7. EMPIRICAL RESULTS

Ten daily international stock market indices were analyzed: the S&P500 (U.S.); FTSE 100 (U.K.); CAC 40 (France); Dax 30 (Germany); Milan MIBTel Index (Italy); To SE 300 (Canada); AORD All ordinaries index (Austr Nikkei 225 Index (Japan); TSEC weighted index (Taiwan the HANG SENG Index (Hong Kong). Data from Janua 2001 to January 19, 2007 were obtained from Datastrea ternational. The percentage log return series were gene taking logarithmic differences of the daily price index, th $y_t = (\ln(p_t) - \ln(p_{t-1})) \times 100$, where p_t is the closing index on day t.

The full data period is divided into a learning sample: ary 1, 2001 to January 10, 2005; and a forecast sample: th trading days from January 11, 2005 to early to mid-Jan 2007. Small differences in end-dates across markets occ due to different market-specific nontrading days. Table 2 s summary statistics from the full sample of the percentag returns of the market indices including sample mean, vari skewness, and kurtosis. All 10 series display the standard erties of daily asset returns: they are heavy tailed and m negatively skewed.

Parameter estimates from the T-CAViaR model are not s to save space. However, for each dataset, the MCMC bu sample size is 15,000 iterations, followed by a samplin riod of 25,000 iterations. To assess mixing and converg the MCMC method is run from five different, randomly erated, starting positions, for each market, at $\alpha = 0.01$, The starting values of each of $(\beta_1, \ldots, \beta_6)$ were chosen on either side of the estimated posterior mean. Converger the same posterior distribution is clear in all five runs for parameter, in each case well before the end of the burn-in ple, in each market. Gelman's R statistics (see Gelman 2005, p. 296) over these five runs, for the Australian m data, were between 1.002 and 1.040 over the six param a typical result; all highlighting fast mixing and clear and cient convergence for the proposed sampling scheme. Fu observed MH acceptance rates in these MCMC runs wer tween 15% and 40%, which Gelman, Roberts, and Gilks (note "yields at least 80% of the maximum efficiency o able" for a Metropolis method. These figures are quite r sentative of the R values and observed MH acceptance during the sampling period across all markets.

Table 2. Summary statistics: percentage log returns on 10 market indices

	Japan	France	Germany	Italy	U.K.
Mean	0.016	−0.0033	0.0027	0.0034	0.0014
Variance	1.955	2.05	2.692	1.226	1.2903
Skewness	−0.080	−0.0411	−0.0993	−0.384	−0.223
	(0.21)	(0.51)	(0.11)	(<0.001)	(0.0003)
Excess kurtosis	1.518	3.504	3.103	5.069	3.946
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
	Canada	**U.S.**	**Taiwan**	**Hong Kong**	**Australia**
Mean	0.025	0.0062	0.0309	0.0201	0.037
Variance	0.760	1.156	2.109	1.3540	0.443
Skewness	−0.465	0.162	−0.018	−0.313	−0.566
	(<0.001)	(0.010)	(0.78)	(<0.001)	(<0.001)
Excess kurtosis	3.842	2.864	1.867	3.772	3.796
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)

NOTE: Percentage log returns are analyzed. p-values based on asymptotic normality are listed, under a null hypothesis of 0 in each case.

7.1 Forecasting VaR Study

VaR is forecast one day ahead for each day in the forecast sample of 500 returns, using a range of competing models from all three quantile estimation classes: nonparametric, semiparametric, and fully parametric. Many financial institutions use historical simulation to forecast VaR; we follow their approach and employ a short-term (ST, last 25 days) and a long-term (LT, last 100 days) percentile: these are nonparametric estimates. We compare two full T-CAViaR models: one with local return threshold (TCAV), the other with US market return threshold (TCAVx), with the two nested versions: AS and SAV, as well as a range of popular GARCH specifications, including GARCH-*n*, GARCH-*t*; GJR-GARCH (GJR), and IGARCH. Except for the RiskMetrics (RMetrics) model, where the parameter is set to 0.94, all estimation uses MCMC methods. Details for the MCMC algorithms can be found in Chen, Gerlach, and So (2006). We also consider the MCMC method of Geweke and Keane (2007) (GK), that employs a smooth mixture of Gaussian regression models. This method requires a long presample period of returns: we use an extra 1000 daily returns, pre-2001, and the same settings as in Geweke and Keane (2007). For each day y_{n+t} in the forecast sample, parameters are estimated for each semiparametric and parametric model, employing the entire previous data set $(y_1, \dots, y_{n+t-1})$ as observations, then forecasts are generated for the next day's α-level quantile. Thus 25,000 (post burn-in) MCMC VaR forecasts are generated each day for each model. The posterior mean quantile forecast for each day is the average of these MCMC forecasts. Each MCMC run takes just under one minute on a standard modern desktop PC, so estimating two years of forecasts (500 days) takes approximately 8 hours, for each model.

Table 3 shows the ratios of the observed VRate to the true nominal level, for $\alpha = 0.01, 0.05$, across all 12 models/methods and 10 markets. The best model's ratio in each market is boxed, while bolding indicates the UC test rejects the model's VRate, at the 5% level. The results are quite different from $\alpha = 0.01$ to 0.05. First, at $\alpha = 0.01$, in all markets, except France, one of the CAViaR models ranks (at least equal) first, with VRate/α closest to 1. Further, the long and short-term HS methods, together with the full set of GARCH models, all have ratios mostly above 1 across the 10 markets. Thus, this group of models consistently under-estimates 1% risk levels in these markets. The GK method is highly variable in its VRate accuracy across the markets, at this level.

A different story applies at $\alpha = 0.05$. Here the models are much closer in performance, with first place rankings spread across models over the 10 markets, and ratios closer to 1 (in location and spread) across most models, compared to $\alpha = 0.01$. However, again both the GK method and the short-term percentile method (ST) consistently under-estimate risk.

Figures 1 and 2 illustrate some VaR forecasts for the Italian MIBTel index returns at $\alpha = 0.01, 0.05$, respectively. For

Table 3. Ratio of VRate/α at $\alpha = 0.01,\ 0.05$ for each model across the 10 markets

Model	Japan	France	Germany	Italy	U.K.	Canada	U.S.	Taiwan	HK	AU
					$\alpha = 0.01$					
ST	**4.6**	**4.2**	**4.8**	**3.8**	**4.6**	**5**	**4**	**4.4**	**3.4**	**4.6**
LT	1.2	1.6	1.8	**2.4**	1.6	1.8	1.6	1.4	1.4	**2**
GARCH-*n*	1.4	1.2	1.6	**3.0**	1.2	**2.4**	1.4	1.6	1.6	**2.8**
GARCH-*t*	1.4	$\boxed{1}$	1.6	**3.0**	1.2	**2**	$\boxed{1}$	1.2	1.4	1.8
GJR	1.2	$\boxed{1}$	1.8	**3.2**	$\boxed{1}$	1.6	$\boxed{1}$	1.4	1.6	**2**
IGARCH	1.6	1.6	**2**	**3.2**	**2.2**	**2.6**	1.8	1.8	1.6	**3**
RMetrics	1.6	1.8	1.6	**3.0**	**2.2**	**2.6**	1.6	**2**	1.8	**2.6**
GK	**0.2**	**0.2**	**0.2**	**3.0**	**2.0**	**5.0**	0.8	$\boxed{1}$	0.8	**9.2**
SAV	$\boxed{0.8}$	0.4	0.6	$\boxed{1}$	$\boxed{1}$	1.8	1.2	$\boxed{1}$	$\boxed{1}$	$\boxed{1.4}$
AS	$\boxed{0.8}$	0.6	$\boxed{0.8}$	0.8	$\boxed{1}$	$\boxed{1.4}$	1.6	0.6	1.2	1.8
TCAV	$\boxed{0.8}$	0.8	1.2	0.8	$\boxed{1}$	1.6	$\boxed{1}$	$\boxed{1}$	0.8	$\boxed{1.4}$
TCAVx	1.4	1.4	1.6	0.8	1.6	**2.2**	–	$\boxed{1}$	0.4	**3.8**
					$\alpha = 0.05$					
ST	1.32	1.4	**1.48**	1.4	**1.56**	**1.52**	1.2	$\boxed{1.44}$	1.36	**1.52**
LT	1.04	1.08	1.16	1.24	1.08	1.4	$\boxed{1.04}$	1.2	1.08	1.28
GARCH-*n*	0.96	$\boxed{1.04}$	1.2	1.08	1.08	1.32	0.84	0.84	1.16	1.32
GARCH-*t*	$\boxed{1}$	1.16	1.24	1.2	1.08	1.32	0.88	0.88	1.2	1.32
GJR	0.96	1.16	1.16	$\boxed{1}$	$\boxed{1}$	1.32	0.84	0.72	$\boxed{1}$	1.24
IGARCH	$\boxed{1}$	1.16	1.12	1.16	1.12	1.24	0.92	$\boxed{1}$	1.2	1.2
RMetrics	0.92	1.12	1.12	1.16	1.04	1.24	0.92	$\boxed{1}$	1.24	1.16
GK	**0.36**	**0.4**	**0.24**	1.36	1.12	**2.2**	**0.6**	**0.48**	**0.6**	**3**
SAV	0.76	$\boxed{1.04}$	$\boxed{1}$	$\boxed{1}$	0.88	1.12	0.72	0.68	0.84	$\boxed{1.08}$
AS	0.8	1.32	1.12	1.12	1.16	$\boxed{1.04}$	**0.64**	**0.6**	**0.6**	1.24
TCAV	0.84	1.32	0.96	1.2	1.12	1.12	**0.6**	**0.56**	1.08	1.36
TCAVx	0.92	1.2	0.96	1.2	1.04	1.44	–	0.76	0.92	**1.76**

NOTE: Boxes indicate closest to 1 in that market, bold indicates the model is rejected by the unconditional coverage test (at a 5% level), for each market. AU: Australia.

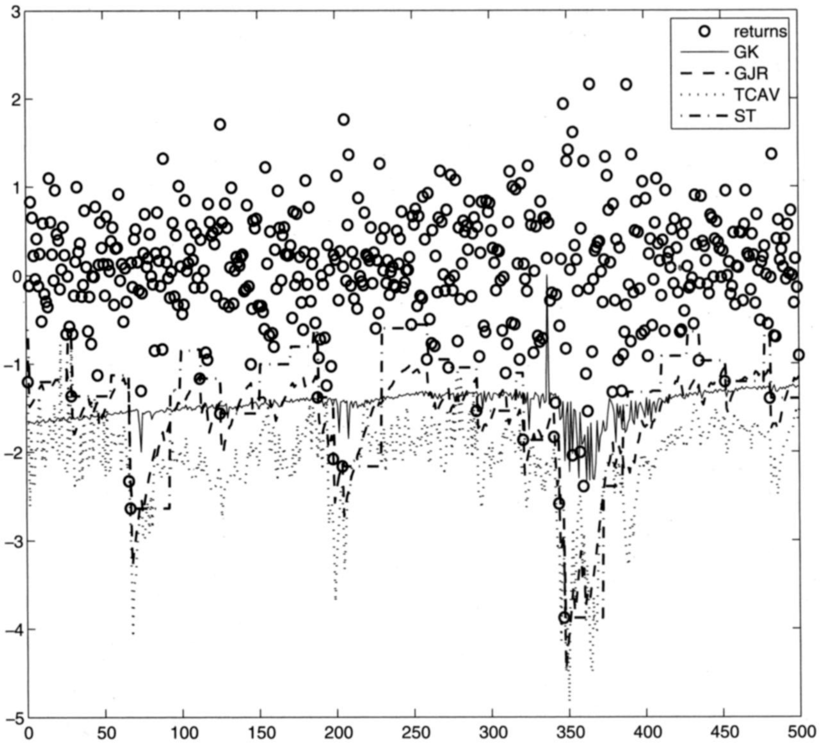

Figure 1. Italian MIBTel index returns January 2005 to January 2007 (circles), together with four sets of forecasted VaR series at $\alpha = 0.01$. Series are GK, GJR, TCAV, and ST.

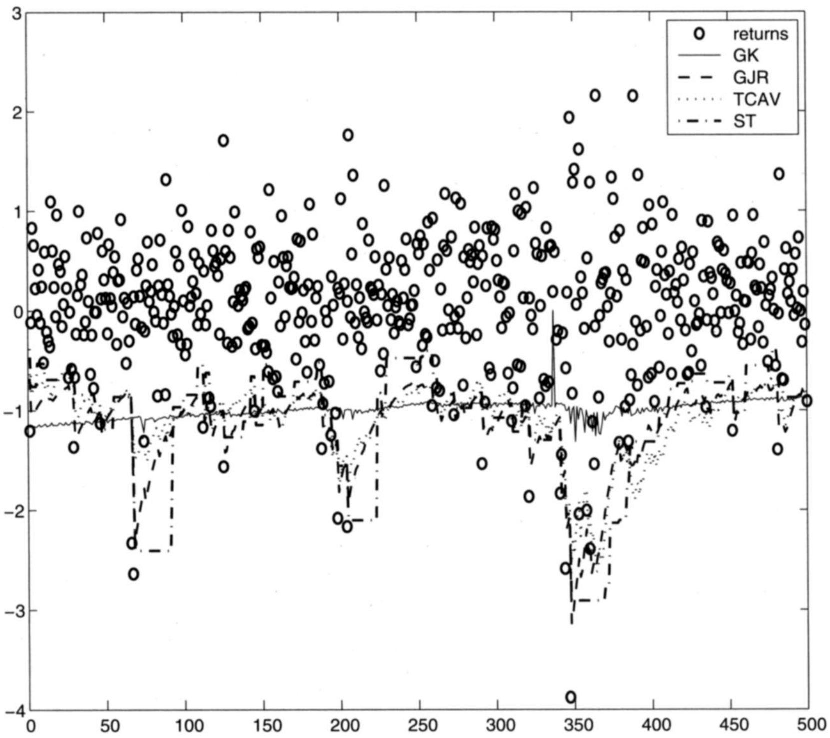

Figure 2. Italian MIBTel index returns January 2005 to January 2007 (circles), together with four sets of forecasted VaR series at $\alpha = 0.05$. Series are GK, GJR, TCAV, and ST.

Italy, at $\alpha = 0.01$ the four CAViaR models ranked first-fourth in VRate/0.01 with 1 (SAV) and 0.8 (AS, TCAV, TCAVx), followed by the GARCH and GK models with ratios, all significantly different to 1, of 3 (GJR and GK) or 3.2; while the ST method had a ratio of 3.8. Figure 1 highlights that the T-CAViaR model was clearly more extreme (and accurate) in its quantile forecasts in this market. The GARCH models' VaR forecasts, similar to the GJR shown, are consistently not large enough in magnitude for a 1% quantile level, compared to the actual returns being forecast. The GK method, while demonstrating that it can capture heteroscedasticity, seemed too smooth in its dynamic quantiles for this market, and did not produce VaR forecasts that reacted to high volatility periods as much as the other models/methods did, its quantile forecasts being less extreme than all other methods in highly volatile return periods and being generally much smoother. The ST method is clearly not optimal for this market either, partly since 25 days is not sufficient to estimate a 1% sample quantile. ST's forecasts are far less smooth than all other models in this case.

Figure 2 shows the VaR forecasts for $\alpha = 0.05$, again for Italy, and for the same four methods. Now we see that the ST method, at least in low-medium volatility periods, is approximating the GARCH (all similar to the GJR shown) models' forecasts, which are all quite similar to the T-CAViaR model. However, the ST method "recovers" slowest after extreme returns or high volatility periods, since any extreme return will affect the sample percentile for exactly 25 days. The GK method's VaR forecasts again seem quite different to the other methods: it is smoother and less reactive to extreme returns. Though it can clearly capture some types of heteroscedasticity, it struggled to accurately capture the specific dynamic volatility process in this data set, at both $\alpha = 0.05, 0.01$.

Table 4 displays summary statistics for the observed VRate/α ratios for each model across the 10 markets, using the numbers in Table 3. The 'Std.' labeled column shows the standard deviation from the expected ratio of 1 (not the sample mean), while '1st' counts the markets where that model had VRate ratio closest to 1 and 'In top 3' counts the markets where the model ranked in the top 3 models by VRate ratio. The results confirm and add to the conclusions above. At $\alpha = 0.01$ the SAV, AS and TCAV (self-exciting) models are most favored across all criteria. In particular, their ratios VRate/α are very close to 1, both in location and spread, and they mostly rank in the top three models for each market. The GARCH models considered all under-estimate risk, with ratios above 1 in most markets, as also do the GK and HS methods. The TCAV model has the lowest deviation in ratios away from 1, average ratio second closest to 1, and is ranked in the top 3 models for all 10 markets. The SAV ranks first in six markets, has average ratio closest to 1, had second lowest deviation in ratios and finished in the top 3 models in seven markets.

Table 4 shows that all the models at $\alpha = 0.05$ have ratios that average close to 1 across markets, except the short-term 25 day percentile (ST), with small deviations away from 1 (except the GK method). The GARCH and CAViaR models are quite comparable at this risk level, and are hard to distinguish between: the TCAV model has average ratio closest to 1, followed by the AS and GJR models. However, again the SAV has the most firs rankings, with four markets, and, together with RMetrics, the most top three rankings, with five markets. The fully threshold CAViaR models both have larger deviations in ratios than the RMetrics, IGARCH, GARCH, ST, LT, AS, and SAV models. The RMetrics and IGARCH models have the lowest deviation in ratios from 1, closely followed by the SAV model. If a conservative risk model is preferred, the SAV is the best candidate, with mean and median ratios at 0.9 and third lowest deviation in ratio from 1.

To help distinguish between the better models at each quantile level, Table 5 shows summary statistics for each model's rank in terms of how close its VRate/α ratio is to 1, across the markets. For ratios that are equidistant from 1, conservative ratios (less than 1) are preferred. Table 5 displays the average, median, standard deviation (from 1) and range of the forecast ranks for each model over the ten markets. At $\alpha = 1\%$, the self-exciting T-CAViaR model has by far the lowest mean rank, equal lowest median rank, and by far the smallest deviation in ranks, away from 1, across the models. In fact, the

Table 4. Summary statistics for VRate/α at $\alpha = 0.01, 0.05$ for each model across the 10 markets

	$\alpha = 1\%$					$\alpha = 5\%$				
	Mean	Median	Std.	1st	In top 3	Mean	Median	Std.	1st	In top 3
ST	**4.34**	**4.5**	**12.64**	0	0	**1.42**	**1.4**	0.21	0	0
LT	1.72	1.7	0.68	0	0	1.16	1.1	0.04	0	2
GARCH-n	1.82	1.6	1.19	0	0	1.08	1.1	0.04	1	3
GARCH-t	1.56	1.4	0.71	2	3	1.13	1.2	0.04	0	2
GJR	1.58	1.5	0.82	3	4	1.04	$\boxed{1.0}$	0.04	3	4
IGARCH	2.14	1.9	1.80	0	0	1.11	1.1	$\boxed{0.02}$	2	4
RMetrics	2.08	1.9	1.55	0	0	1.09	1.1	$\boxed{0.02}$	1	$\boxed{5}$
GK	2.24	0.9	10.03	1	2	1.04	0.6	**0.84**	0	0
SAV	$\boxed{1.02}$	$\boxed{1.0}$	0.16	$\boxed{6}$	7	0.91	0.9	0.03	$\boxed{4}$	$\boxed{5}$
AS	1.04	0.9	0.16	4	6	0.96	1.1	0.08	1	1
TCAV	1.04	$\boxed{1.0}$	$\boxed{0.08}$	5	$\boxed{10}$	$\boxed{1.02}$	1.1	0.08	0	3
TCAVx	1.58	1.4	1.34	1	1	1.13	$\boxed{1.0}$	0.12	0	3

NOTE: Boxes indicate the favored model, bold indicates the least favoured model, in each column; 'Std.' is the standard deviation in ratios from an expected value of 1; '1st' indicates the number of markets where that model's VRate ratio ranked closest to 1; 'In top 3' counts the number of markets where the model's VRate ratio ranked in the top 3 models.

Table 5. Summary statistics for model ranks, in terms of VRate/α, at $\alpha = 0.01, 0.05$ across the 10 markets

	$\alpha = 1\%$				$\alpha = 5\%$			
	Mean	Median	Std.	Range	Mean	Median	Std.	Range
ST	**11.75**	**12.0**	**128.58**	1	**10.50**	11.0	**102.33**	5
LT	7.10	7.5	45.89	6	6.40	5.5	43.06	11
GARCH-*n*	6.90	7.3	41.00	5	5.20	5.3	24.78	7.5
GARCH-*t*	4.85	5.3	20.36	6	5.85	6.0	32.86	8.5
GJR	5.40	5.0	32.44	9	4.35	4.8	18.81	6.5
IGARCH	9.65	9.8	84.03	3	4.50	4.8	19.72	7.5
RMetrics	9.00	9.5	74.72	5.5	$\boxed{4.20}$	$\boxed{3.8}$	$\boxed{16.94}$	7.5
GK	7.27	9.0	57.04	9.5	10.27	**11.8**	96.49	3
SAV	3.00	$\boxed{2.5}$	8.22	6	4.55	$\boxed{3.8}$	25.97	9
AS	3.45	3.3	10.42	5.5	7.25	8.8	55.08	10.5
TCAV	$\boxed{2.35}$	$\boxed{2.5}$	$\boxed{2.25}$	1.5	7.15	8.3	51.58	8
TCAVx	6.06	6.0	33.61	7.5	6.17	6.5	40.47	9

NOTE: Boxes indicate the favored model, bold indicates the least favored model, in each column; 'Std.' is the standard deviation in ranks from the value of 1.

TCAV model ranks in the top 3 in every market, and thus has the smallest range in ranks (except for ST which only ranked 11th or 12th). The SAV and then AS CAViaR models are next best, with SAV having the equal best median rank and 2nd best mean rank and 2nd lowest deviation in ranks, and the AS model finishing 3rd on mean, median, and deviation in rank measures. The GARCH-*n*, GARCH-*t*, GJR-GARCH, and TCAVx models typically ranked just behind the TCAV, SAV, and AS CAViaR models in 4th–7th placing in each market, followed by the long-term percentile (LT) and GK methods; while RMetrics, IGARCH and short-term (ST) percentile methods typically ranked in the last three places in each market. Clearly the TCAV model has forecast dynamic quantiles accurately and comparatively the best over the twelve models and 10 markets considered, at $\alpha = 0.01$. Further, three of the CAViaR models do best as a group, followed by the stationary GARCH specifications, the GARCH-*t*'s fat-tails outperforming the GARCH-n as expected, then the TCAVx, GK and remaining methods. RiskMetrics and IGARCH do not forecast VaR well at this quantile level at all, nor did the HS methods.

For $\alpha = 0.05$, the RiskMetrics method now has the lowest average and equal lowest median rank across markets, closely followed by the GJR-GARCH, IGARCH, and SAV models in mean rank. The RMetrics method also has the lowest deviation in ranks away from 1. Clearly, no model dominated in rank, and thus risk ratio accuracy, at this risk level, since the summary rank measures are above 4 in average and have much larger deviation, compared to those for $\alpha = 0.01$. The SAV model finished with the equal best median rank and 4th best in both mean and standard deviation in ranks across markets. These four models are hard to separate at this quantile level, but the RiskMetrics method is marginally favored considering both Tables 4 and 5. The AS and T-CAViaR models now typically rank in the middle (i.e., 6th to 9th) of the 12 models.

To summarize, the fully nonlinear self-exciting T-CAViaR model is favored for accurate dynamic quantile forecasting when $\alpha = 0.01$, ranking first in all summary rank measures and first in median and deviation in VRate ratios from 1; followed by the SAV and AS CAViaR models. Further, at $\alpha = 0.05$, four models: RMetrics, GJR, IGARCH, and SAV, are hard to separate, though RMetrics could be favoured since it ranked first in all summary rank measures and had equal lowest deviation in VRate ratios. We now formally test these models, as a final point of comparison.

Table 6 counts the number of rejections for each model, over the 10 markets, at the 5% significance level, for each of the three tests considered: UC, CC, and the DQ tests. Four lags were used, as in Engle and Manganelli (2004). At $\alpha = 0.01$ most of the models are rejected in most of the markets, mainly by the DQ test. The CAViaR models fare best as a group, with the self-exciting T-CAViaR model rejected in the least number of markets for all three tests, and for all tests combined (four rejections across 10 markets), followed by AS with five, TCAVx with six, and SAV with seven rejections. The GK method is only rejected in six markets by DQ, but in nine markets overall, while all the GARCH models are rejected in at least eight markets; ST, RMetrics, and IGARCH are rejected in all 10 markets. At $\alpha = 0.05$ the HS (ST, LT) and GK methods are rejected in seven or nine markets, while the CAViaR models are rejected in two or three markets each; the GARCH and RMetrics models fare best with only 1 or 2 rejections across markets.

In summary, the CAViaR family of models are highly competitive at worst, and far more accurate at best, at dynamic quantile VaR forecasting, compared to a range of popular and well-known VaR methods. At $\alpha = 0.05$ most of the models forecasted similarly with no clear standout, though RiskMetrics and the CAViaR SAV model finished best in mean and median ranking, respectively, in terms of violation rates, across markets; and the RiskMetrics, GJR and GARCH-*n* models are rejected the least across all VaR forecast methods. The full T-CAViaR model, though still performing well in overall violation rate across markets, is perhaps an unnecessarily complex model at quantile level 0.05, the simpler symmetric models fared better at this level. However, the self-exciting T-CAViaR model performed in a comparably superior fashion to the other models, followed by the remaining CAViaR specifications, for $\alpha = 0.01$. Its VaR forecasts display violation rates that are consistently closest to nominal and the model is rejected the least number of times across the 10 markets. The RiskMetrics, IGARCH, stationary GARCH, GK, and HS methods are simply not competitive with the CAViaR models at forecasting risk levels or

Table 6. Counts of model rejections for three standard quantile coverage tests, across the 10 markets

	$\alpha = 1\%$				$\alpha = 5\%$			
Model	UC p	CC p	DQ4 p	Total	UC p	CC p	DQ4 p	Total
ST	**10**	**10**	**10**	**10**	5	5	**9**	**9**
LT	2	2	9	9	$\boxed{0}$	2	7	7
GARCH-n	3	3	9	9	$\boxed{0}$	$\boxed{0}$	$\boxed{1}$	$\boxed{1}$
GARCH-t	2	1	9	9	$\boxed{0}$	$\boxed{0}$	2	2
GJR	2	1	8	8	$\boxed{0}$	$\boxed{0}$	$\boxed{1}$	$\boxed{1}$
IGARCH	5	4	**10**	**10**	$\boxed{0}$	$\boxed{0}$	$\boxed{1}$	$\boxed{1}$
RMetrics	5	4	**10**	**10**	$\boxed{0}$	$\boxed{0}$	$\boxed{1}$	$\boxed{1}$
GK	7	4	6	9	**8**	**7**	8	**9**
SAV	$\boxed{0}$	$\boxed{0}$	7	7	$\boxed{0}$	$\boxed{0}$	2	2
AS	$\boxed{0}$	$\boxed{0}$	5	5	3	$\boxed{0}$	$\boxed{1}$	3
TCAV	$\boxed{0}$	$\boxed{0}$	$\boxed{4}$	$\boxed{4}$	2	1	$\boxed{1}$	3
TCAVx	2	$\boxed{0}$	6	6	$\boxed{0}$	$\boxed{0}$	2	2

NOTE: Boxes indicate the favored model, bold indicates the least favored model, in each column.

dynamics, at this more extreme quantile: all over-estimated the risk levels and/or do not effectively capture the dynamics of risk and are subsequently rejected by the formal tests, especially the DQ test, in most markets at $\alpha = 0.01$.

8. CONCLUSION

This article proposes a new fully nonlinear CAViaR model for dynamic quantile estimation. Bayesian MCMC methods are adapted to this family, employing the well-known link between the quantile estimation criterion function and the Skewed-Laplace density. MCMC estimation proves favourably precise and efficient, in simulations from a GARCH model, compared to the classical numerically optimized quantile criterion function, at the quantile levels 0.05 and 0.01. A VaR forecasting study reveals that the fully nonlinear self-exciting T-CAViaR model compares most favorably in terms of violation rates and independence of violations, to more parsimonious CAViaR models at level 0.01, while the simplest CAViaR SAV model fares better than the nonlinear CAViaR models at level 0.05. A range of well-known GARCH models, including RiskMetrics, historical simulation and a semiparametric smoothly mixing regression approach are not competitive at the quantile level 0.01, across the 10 markets. However, at level 0.05, most models forecast reasonably similarly, with results marginally favoring simpler specifications such as RiskMetrics, GJR-GARCH, and the CAViaR SAV specification.

ACKNOWLEDGMENTS

We thank an AE and an anonymous referee for their insightful and helpful comments and suggestions which improved the article. Cathy Chen is supported by the grant 96-2118-M-002-MY3 from the National Science Council (NSC) of Taiwan. The revision of this article was conducted while Richard Gerlach was a Visiting Scholar at Feng Chia University supported by the Mathematics Research Promotion Center, NSC, Taiwan.

[Received July 2008. Revised January 2010.]

REFERENCES

Berkowitz, J., Christoffersen, P., and Pelletier, D. (2010), "Evaluating Value-at-Risk Models With Desk-Level Data," *Management Science*, to appear. [486]

Black, F. (1976), "Studies in Stock Price Volatility Changes," in *Proceedings of the Business and Economic Statistics Section*, Alexandria, VA: American Statistical Association, pp. 177–181. [481,483]

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327. [481]

Brooks, C. (2001), "A Double-Threshold GARCH Model for the French Franc/Deutschmark Exchange Rate," *Journal of Forecasting*, 20, 135–143. [481,483]

Chen, C. W. S., and So, M. K. P. (2006), "On a Threshold Heteroscedastic Model," *International Journal of Forecasting*, 22, 73–89. [481,484]

Chen, C. W. S., Gerlach, R., and So, M. K. P. (2006), "Comparison of Non-Nested Asymmetric Heteroskedastic Models," *Computational Statistics and Data Analysis*, 51 (4), 2164–2178. [487]

Christoffersen, P. (1998), "Evaluating Interval Forecasts," *International Economic Review*, 39, 841–862. [486]

Engle, R. F. (1982), "Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of United Kingdom Inflations," *Econometrica*, 50, 987–1007. [481]

Engle, R. F., and Manganelli, S. (2004), "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," *Journal of Business & Economic Statistics*, 22, 367–381. [481,483,486,490]

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2005), *Bayesian Data Analysis* (2nd ed.), Boca Raton, FL: Chapman & Hall. [484,486]

Gelman, A., Roberts, G. O., and Gilks, W. R. (1996), "Efficient Metropolis Jumping Rules," in *Bayesian Statistics 5*, Oxford: Clarendon, pp. 599–608. [484,486]

Geraci, M., and Bottai, M. (2007), "Quantile Regression for Longitudinal Data Using the Asymmetric Laplace Distribution," *Biostatistics*, 8, 140–154. [481]

Geweke, J., and Keane, M. (2007), "Smoothly Mixing Regressions," *Journal of Econometrics*, 138, 252–290. [481,487]

Giacomini, R., and Komunjer, I. (2005), "Evaluation and Combination of Conditional Quantile Forecasts," *Journal of Business & Economic Statistics*, 23, 416–431. [481]

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 Stock," *Journal of Finance*, 48, 1779–1801. [483]

Hastings, W. K. (1970), "Monte-Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, 57, 97–109. [484]

Jorion, P. (1996), "Risk: Measuring the Risk in Value at Risk," *Financial Analysis Journal*, 52, 47–56. [481,482]

Koenker, R., and Bassett, G. (1978), "Regression Quantiles," *Econometrica*, 46, 33–50. [481]

Kuester, K., Mittnik, S., and Paolella, M. S. (2006), "Value-at-Risk Prediction: A Comparison of Alternative Strategies," *Journal of Financial Econometrics*, 4, 53–89. [481]

Kupiec, P. H. (1995), "Techniques for Verifying the Accuracy of Risk Measurement Models," *Journal of Derivatives*, 3, 73–84. [486]

Li, C. W., and Li, W. K. (1996), "On a Double-Threshold Autoregressive Heteroscedastic Time Series Model," *Journal of Applied Econometrics*, 11, 253–274. [483]

Manganelli, S., and Engle, R. F. (2004), "A Comparison of Value-at-Risk Models in Finance," in *Risk Measures for the 21st Century*, ed. G. Szegö, Chichester: Wiley. [481,482]

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953), "Equation of State Calculations by Fast Computing Machines," *Journal of Chemical Physics*, 21, 1087–1092. [484]

Morgan, J. P. (1996), *RiskMetrics—Technical Document* (4th ed.), New York: Morgan Guaranty Trust Company. [482]

Nelson, D. B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach Econometrica," *Econometrica*, 59, 347–370. [483]

Tong, H. (1978), "On a Threshold Model" in *Pattern Recognition and Signal Processing*, ed. C. H. Chen, Amsterdam: Sijhoff & Noordhoff. [483]

——— (1983), *Threshold Models in Non-Linear Time Series Analysis. Lecture Notes in Statistics*, Vol. 21, New York: Springer-Verlag. [483]

Tsay, R. S. (2005), *Analysis of Financial Time Series* (2nd ed.), Hoboken: Wiley. [482]

Tsionas, E. G. (2003), "Bayesian Quantile Inference," *Journal of Statistical Computation and Simulation*, 9, 659–674. [481,482]

Wong, C. M., and So, K. P. (2003), "On Conditional Moments of GARCH Models, With Applications to Multiple Period Value at Risk Estimation," *Statistica Sinica*, 13, 1015–1044. [486]

Yu, K., and Moyeed, R. A. (2001), "Bayesian Quantile Regression," *Statistics and Probability Letters*, 54, 437–447. [481-483]

Yu, P. L. H., Li, W. K., and Jin, S. (2010), "On Some Models for Value-at-Risk," *Econometric Reviews*, to appear. [483]

Zakoian, J. (1994), "Threshold Heteroskedastic Model," *Journal of Economic Dynamics and Control*, 18, 931–955. [481,483]