International Journal of Forecasting 28 (2012) 557–574

Contents lists available at SciVerse ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting Value-at-Risk using nonlinear regression quantiles and the intra-day range

Cathy W.S. Chen [a,*], Richard Gerlach [b], Bruce B.K. Hwang [a], Michael McAleer [c,d,e]

[a] *Graduate Institute of Statistics and Actuarial Science, Feng Chia University, Taiwan*
[b] *The University of Sydney Business School, Australia*
[c] *Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands*
[d] *Institute of Economic Research, Kyoto University, Japan*
[e] *Department of Quantitative Economics, Complutense University of Madrid, Spain*

ARTICLE INFO

Keywords:
Value-at-Risk
CAViaR model
Skewed–Laplace distribution
Intra-day range
Backtesting
Markov chain Monte Carlo

ABSTRACT

Some novel nonlinear threshold conditional autoregressive VaR (CAViaR) models are proposed that incorporate intra-day price ranges. Model estimation is performed using a Bayesian approach via the link with the Skewed–Laplace distribution. The performances of a range of risk models during the 2008–09 financial crisis are examined, including an evaluation of the way in which the crisis affected the performance of VaR forecasting. An empirical analysis is conducted on five Asia-Pacific Economic Cooperation stock market indices and two exchange rate series. Standard back-testing criteria are used to measure and assess the forecast performances of a variety of risk models. The proposed threshold CAViaR model, incorporating range information, is shown to forecast VaR more effectively and more accurately than other models, across the series considered.

1. Introduction

The global financial crisis (GFC) of 2008–09 has (once again) called into question financial risk management practices, and one issue is whether risk measures can actually be forecast accurately enough for the purpose. This paper adds to this area by proposing and assessing some novel univariate, semi-parametric range-based conditional autoregressive VaR (CAViaR) models. The aim is to generate more accurate forecasts of VaR for univariate asset returns, for single fixed-weight portfolio returns and for exchange rate return series, to help achieve better risk measurement and risk management practice. We attempt this by incorporating intra-day high-low price range data into the CAViaR model. We then examine whether this proposal adds to the accuracy of VaR forecasts during the GFC; the evidence presented here suggests that such is indeed the case.

Quantitative risk measure forecasting has become very important, at least since the market crash in 1987, and even more so after the recent GFC. The Basel II Accord is designed to monitor and encourage sensible risk-taking, using appropriate models to calculate VaR and daily capital charges. VaR is now a standard tool in risk management. It was pioneered by the Morgan Corporation in 1993 via their RiskMetrics system, and has been more formally defined by Jorion (1996) as an estimate of the probability and size of the worst potential or expected loss over a given time horizon with a specified probability. Mathematically:

$$P_r\left(\Delta V(l) \le -\mathrm{VaR}|\mathcal{F}_{t-1}\right) = \alpha,$$

where $\Delta V(l)$ is the change in the asset value over time period l, α is the probability level, and $\mathcal{F}_{t-1}$ denotes the information set at time $t-1$. Finding models and methods for VaR forecasting is an on-going challenge for financial practitioners and statisticians.

* Corresponding author. Tel.: +886 4 24517250x4412; fax: +886 4 2451 7092.
E-mail address: chenws@mail.fcu.edu.tw (C.W.S. Chen).

0169-2070/$ – see front matter © 2012 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.ijforecast.2011.12.004

The use of intra-day high-low price range data has been known to increase the efficiency of volatility estimation, compared to the use of simple daily returns data, since at least Garman and Klass (1980) and Parkinson (1980). In this paper, we propose a new semi-parametric family of CAViaR models that incorporate the use of range information, rather than returns data, as was utilised in the CAViaR models of Engle and Manganelli (2004). We therefore examine the question of whether the efficiency gains of range data over return data for volatility estimation also apply to the estimation and forecasting of Value-at-Risk.

We adapt the Bayesian estimation methods of Yu and Moyeed (2001), applying the link between quantile estimation and the Skewed–Laplace distribution, first discussed by Koenker and Machado (1999), to the models in the extended CAViaR family in a systematic way. Further, we conduct a comparison of these methods with the frequentist estimation of Engle and Manganelli (2004), in regard to the forecasts produced from these models.

In this paper we discuss the selection of optimal risk models; examine the performances of risk management strategies during the 2008–09 GFC; evaluate the way in which the crisis affected risk management practices, forecasts of VaR and daily capital charges; and discuss the diagnostic checking of VaR methods. We conduct a "horse-race" of our proposed range-based CAViaR models with some existing, competing models, during the GFC period, for some individual market returns, a portfolio of these market returns, and two exchange rate return series.

The paper is structured as follows. In Section 2, methods for VaR forecasting are reviewed and the new CAViaR specifications are presented. Section 3 discusses the estimation of VaR models and criteria for measuring VaR performance. An empirical analysis is conducted in Section 4 on five Asia-Pacific Economic Cooperation (APEC) stock market indices, namely Standard and Poor's 500 Index, Nikkei 225, TAIEX, HSI and KOSPI, for forecasting VaR from August 2008 to April 2010. Finally, some concluding remarks are given in Section 5.

2. Value-at-Risk—models and methods

A single period VaR is proportional to the conditional quantile of the single period return distribution. Mathematically, for a given value α, $0 \leq \alpha \leq 1$, the αth quantile of the random variable y is defined as $q_\alpha(y) = \inf\{y|F(y) \geq \alpha\}$, where F is the CDF of y.

2.1. VaR estimation methods

There are many VaR estimation methods in the literature, which can be classified into three broad categories. First, non-parametric VaR estimation, where no, or very few, or non-restrictive, assumptions are made on the distribution of returns, e.g. historical simulation, which uses past sample return quantiles. At the other end of the scale there is fully parametric estimation, where VaR is usually constructed assuming a specific choice of the unconditional and/or conditional return distribution, and also of the model dynamics. One such example is the generalized autoregressive conditional heteroskedastic (GARCH) volatility equations (proposed by Bollerslev, 1986; Engle, 1982), with specific noise distributions, such as Gaussian, Student-t, and skewed Student-t (see Hansen, 1994). Chen, Gerlach, Lin, and Lee (in press) and Kuester, Mittnik, and Paolella (2006) both consider a range of parametric models for forecasting VaR, including standard, threshold nonlinear and Markov switching GARCH specifications (see e.g. Haas, Mittnik, & Paolella, 2006; or Guidolin & Timmermann, 2006), with various probability distributions. Finally, semi-parametric methods often make assumptions about the model dynamics but not the error distribution; e.g., Engle and Manganelli (2004) propose direct dynamic quantile regression (see Koenker & Bassett, 1978) for calculating VaR, denoted CAViaR; while Gerlach, Chen, and Chan (2011) propose a family of nonlinear CAViaR models.

McAleer, Jimenez-Martin, and Perez-Amaral (2010a) consider mixing alternative risk models, and discuss the choice between conservative and aggressive risk management, as well as evaluating the effects of the Basel II Accord for risk management. McAleer and da Veiga (2008b) compare the performance of single-index and portfolio models in forecasting VaR, while McAleer, Jimenez-Martin, and Perez-Amaral (2010b) provide a method for choosing one risk model at the beginning of the period, and then modify the forecast depending on the recent history of violations.

In this paper we use methods from all three of the above-mentioned classes. Historical simulation is employed in the non-parametric category, where we use two sample percentiles: a short-term (ST, the last 25 days) and a long-term (LT, the last 100 days).

2.2. Parametric VaR

The main parametric methods used are the RiskMetrics and GARCH models. In our study, the IGARCH(1, 1) of RiskMetrics with Gaussian errors, as proposed by Morgan (1996), and the GARCH(1, 1) model with Gaussian and Student-t errors, are considered in the empirical analysis; much of the literature on VaR forecasting uses these models as benchmarks. The models are specified as:

Model A: GARCH model

$$y_t = \mu_t + a_t, \quad \mu_t = \phi_0 + \phi_1 y_{t-1},$$

$$a_t = \varepsilon_t \sqrt{h_t}, \quad \text{where } \varepsilon_t \overset{i.i.d}{\sim} D(0, 1),$$

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1}.$$

Model B: RiskMetrics model

$$y_t = a_t$$

$$a_t = \varepsilon_t \sqrt{h_t}, \quad \text{where } \varepsilon_t \overset{i.i.d}{\sim} N(0, 1)$$

$$h_t = (1 - \lambda) a_{t-1}^2 + \lambda h_{t-1}.$$

Under each model, the one-step-ahead VaR at the α% quantile level is computed as $\text{VaR}_t = \mu_t + D_\alpha^{-1}\sqrt{h_t}$, where D^{-1} is the inverse CDF for the distribution D. The parameters of the GARCH models are estimated by Bayesian Markov chain Monte Carlo (MCMC), as discussed by Chen, Chiang, and So (2003) and in the next section.

2.3. Semi-parametric quantile regression and CAViaR

Koenker and Bassett (1978) suggest that, based on a sample of i.i.d. realizations $\{y_t\}$ of y, the quantile $b = q_\alpha(y)$ can be estimated by solving the following minimization problem:

$$\min_{b\in\Re}\left[\sum_t (y_t - b)(\alpha - I\{y_t < b\})\right].$$

Engle and Manganelli (2004) propose some time series models for the quantile, i.e. b becomes b_t and the i.i.d. assumption is relaxed, a method called CAViaR, and use this same criterion to estimate the unknown parameters in the models for b_t. Let y_t be an asset, market, portfolio or exchange rate return at time t, and $\boldsymbol{\beta}_\alpha$ the vector of $q+r$ unknown parameters, $(\beta_1,\dots,\beta_q,\beta_{q+1},\dots,\beta_{q+r})'$, for the α-quantile model. For notational convenience, we let $f_t(\boldsymbol{\beta}) = f_t(y_t,\boldsymbol{\beta}_\alpha)$ denote the time t conditional α level quantile. A general specification of VaR at time t is:

$$f_t(\boldsymbol{\beta}) = \beta_0 + \sum_{i=1}^{q}\beta_i f_{t-i}(\boldsymbol{\beta}) + g(\beta_{q+1},\dots,\beta_{q+r},\mathcal{F}_{t-1}),$$

where $g()$ is a function of a finite number of lagged returns and model parameters, thus linking the α quantile $f_t(\boldsymbol{\beta})$ to past returns, which are a subset of all past information, denoted as $\mathcal{F}_{t-1}$. $\beta_i f_{t-i}(\boldsymbol{\beta})$ is the autoregressive term which ensures smooth quantile changes over time. Three general CAViaR specifications of Engle and Manganelli (2004) are:

(1) Symmetric Absolute Value (SAV):

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

(2) Asymmetric Slope (AS):

$$f_t(\boldsymbol{\beta}) = \beta_1 + \beta_2 f_{t-1}(\boldsymbol{\beta}) + (\beta_3 I_{(y_{t-1}>0)} + \beta_4 I_{(y_{t-1}<0)})|y_{t-1}|. \quad (2)$$

(3) Indirect GARCH(1, 1) (IG):

$$f_t(\boldsymbol{\beta}) = (\beta_1 + \beta_2 f^2_{t-1}(\boldsymbol{\beta}) + \beta_3 y^2_{t-1})^{1/2}. \quad (3)$$

Giacomini and Komunjer (2005) find that CAViaR is most accurate at the 1% quantile level, but the GARCH model with Gaussian distributed errors is better than CAViaR at the 5% quantile level. Gerlach et al. (2011) find similar results across a range of financial market indices.

Yu, Li, and Jin (2010) extend these CAViaR models using two approaches, namely the threshold and mixture type indirect-GARCH CAViaR models. Gerlach et al. (2011) also propose a nonlinear CAViaR model. We adopt the threshold CAViaR (TCAV) model of Gerlach et al. (2011), and the threshold-type indirect-VaR model of Yu et al. (2010) as follows:

(4) Threshold CAViaR (TCAV)

$$f_t(\boldsymbol{\beta}) = \begin{cases}\beta_1 + \beta_2 f_{t-1}(\boldsymbol{\beta}) + \beta_3\,|y_{t-1}|, & z_{t-1}\le\gamma\\ \beta_4 + \beta_5 f_{t-1}(\boldsymbol{\beta}) + \beta_6\,|y_{t-1}|, & z_{t-1}>\gamma,\end{cases} \quad (4)$$

where $\boldsymbol{z}$ is an observed threshold variable, which can be exogenous or self-exciting (i.e. $z_t = y_t$), and γ is the threshold value, typically set as $\gamma = 0$. We extend the model slightly by estimating this parameter in this paper, while Gerlach et al. (2011) fix $\gamma = 0$. Further:

(5) Threshold Indirect GARCH(1, 1) (TIG):

$$f_t(\boldsymbol{\beta}) = \begin{cases}(\beta_1 + \beta_2 f^2_{t-1}(\boldsymbol{\beta}) + \beta_3 y^2_{t-1})^{1/2}, \\ \quad \text{if } y_{t-1}<\gamma,\\ (\beta_4 + \beta_5 f^2_{t-1}(\boldsymbol{\beta}) + \beta_6 y^2_{t-1})^{1/2},\\ \quad \text{if } y_{t-1}\ge\gamma.\end{cases} \quad (5)$$

2.4. Proposed range-based CAViaR models

There are several advantages of using the intra-day high-low price range directly for volatility measurement and forecasting, relative to the use of absolute or squared return data, or intra-day returns. Many papers have shown the intra-day range to be an efficient measure of daily volatility (e.g., see Parkinson, 1980). Mandelbrot (1971) proposes the use of the range for evaluating the existence of a long-term dependence on asset prices; Garman and Klass (1980) show that high-low price-range data contain more information regarding volatility than opening to closing prices. Beckers (1983) uses the range estimator to incorporate past information for different variance measures. Gallant, Hsu, and Tauchen (1999) and Alizadeh, Brandt, and Diebold (2002) incorporate the range into the stochastic volatility model. Brandt and Jones (2006) propose a range-based EGARCH model, using a link between the range and intra-day volatility, and show that their model has a favourable out-of-sample volatility forecasting performance. Chou (2005) proposes the Conditional Autoregressive Range (CARR) model for the high and low range of asset prices. Chen, Gerlach, and Lin (2008) allow the intra-day high and low price range to depend nonlinearly on past information, or on an exogenous variable such as US market information, finding an increased accuracy of volatility estimation using the CARR and GARCH models. Here, we propose a family of CAViaR models that incorporate intra-day price range information.

In the same spirit as Chou (2005) and Chen et al. (2008), we extend the CAViaR models in Eqs. (2), (4) and (5) and incorporate the intra-day high-low price range into the following models:

(6) Range Value (RV):

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

(7) Threshold Range Value (TRV):

$$f_t(\boldsymbol{\beta}) = \begin{cases}\beta_1 + \beta_2 f_{t-1}(\boldsymbol{\beta}) + \beta_3 R_{t-1}, & \text{if } R_{t-1}\le\gamma,\\ \beta_4 + \beta_5 f_{t-1}(\boldsymbol{\beta}) + \beta_6 R_{t-1}, & \text{if } R_{t-1}>\gamma.\end{cases} \quad (7)$$

The first model has the same form as the SAV model in Eq. (1), but replaces the absolute return with the intra-day price range R_{t-1}. The TRV has the same form as the TCAV in Eq. (4), again replacing the return data with the range data. The following model makes the same adjustments to the TIG model in Eq. (5):

(8) Threshold Range Indirect GARCH(1, 1) (TRIG):

$$f_t(\boldsymbol{\beta}) = \begin{cases} (\beta_1 + \beta_2 f_{t-1}^2(\boldsymbol{\beta}) + \beta_3 R_{t-1}^2)^{1/2}, \\ \quad \text{if } R_{t-1} \le \gamma, \\ (\beta_4 + \beta_5 f_{t-1}^2(\boldsymbol{\beta}) + \beta_6 R_{t-1}^2)^{1/2}, \\ \quad \text{if } R_{t-1} > \gamma. \end{cases} \tag{8}$$

Here, R_t is the intra-day range at time t, and γ is the threshold value. The RV model responds symmetrically to the past range, while the TRV and TRIG allow for different responses to high and low ranges.

3. Estimation and forecast evaluation

Using the Koenker and Bassett (1978) regression quantile framework, the unknown parameters of CAViaR models can be estimated by optimising a criterion function. The αth regression quantile is defined as the solution, $\boldsymbol{\beta}_\alpha$, of the criterion function:

$$\min \sum (y_t - f_t(\boldsymbol{\beta})) \left\{\alpha - I_{(-\infty,0)}(y_t - f_t(\boldsymbol{\beta}))\right\}, \tag{9}$$

where $f_t(\boldsymbol{\beta})$ is the model for the αth regression quantile. Based on a sample of data $y_1, \ldots, y_n$, the function in Eq. (9) can be minimised numerically to find $\hat{\boldsymbol{\beta}}_\alpha$, as was done by Engle and Manganelli (2004) for CAViaR models (1)–(3). Gerlach et al. (2011) also use this method for estimation in the TCAV in Eq. (4), and show that, for simulated data, the Bayesian estimates using MCMC are more efficient for that model. We discuss this approach now.

It has recently been shown that the quantile regression criterion function is related to the likelihood function for the skewed-Laplace distribution. This result allows (maximum) likelihood estimation, and has motivated Bayesian solutions for this problem, as proposed by Yu and Moyeed (2001), Tsionas (2003), Yu and Zhang (2005) and Geraci and Bottai (2007), and subsequently extended by Chen et al. (in press). These designs all involve MCMC computational methods, due to the non-standard form of the posterior resulting from the skewed-Laplace likelihood.

3.1. Frequentist estimation

First, we define the data vectors as $\boldsymbol{y} = (y_1, \ldots, y_n)'$ for the asset returns and $\boldsymbol{R} = (R_1, \ldots, R_n)'$ for the intra-day range data. If we assume that the returns follow a skewed-Laplace, i.e. $y_t \overset{i.i.d}{\sim} SL(f_t(\boldsymbol{\beta}), \tau, \alpha)$, then the following density function results:

$$f(y_t; f_t(\boldsymbol{\beta}), \tau, \alpha) = \frac{\alpha(1-\alpha)}{\tau} \exp\left[-\rho_\alpha\left(\frac{y_t - f_t(\boldsymbol{\beta})}{\tau}\right)\right],$$

where $\rho_\alpha(u) = u(\alpha - I(u < 0))$, $f_t(\boldsymbol{\beta})$ is the mode and $\tau > 0$ is a scale parameter. Under this assumption, the likelihood function for any CAViaR model, including Eqs. (1)–(8), is then:

$$L_\alpha(\boldsymbol{\beta}, \tau, \gamma; \boldsymbol{y}, \boldsymbol{R}) \propto \tau^{-n} \exp\left\{ -\tau^{-1} \times \left[\sum_{t=1}^{n} (y_t - f_t(\boldsymbol{\beta}))\left(\alpha - I_{(-\infty,0)}(y_t - f_t(\boldsymbol{\beta}))\right)\right]\right\}. \tag{10}$$

As such, the $\hat{\boldsymbol{\beta}}_\alpha$ that minimises Eq. (9) also maximises Eq. (10). This estimate can then simply be plugged into the formula for $f_{n+1}(\boldsymbol{\beta})$ in order to forecast VaR.

3.2. Bayesian estimation and forecasting

Bayesian inference requires a prior distribution to be specified for the unknown parameters, combined with the likelihood function. Assuming that the parameters $(\boldsymbol{\beta}, \tau, \gamma)$ are a priori independent, we choose $\pi(\tau) \propto \tau^{-1}$, the standard Jeffreys' prior, and $\pi(\boldsymbol{\beta}) \propto 1$, as in Gerlach et al. (2011). When considering two regimes, a flat prior on the threshold limit γ is Unif(u, l), where the (u, l) are chosen as suitable quantiles of the threshold variable, to allow a reasonable sample size in each regime for inference.

MCMC methods sample from the joint posterior distribution of the unknown model parameters for estimation, inference and forecasting. Here, groups of parameters are defined for the following sampling scheme:

$$p(\boldsymbol{\beta}|\boldsymbol{y}, \boldsymbol{R}, \gamma),\ p(\tau|\boldsymbol{y}, \boldsymbol{R}, \boldsymbol{\beta}, \gamma),\ \text{and}\ p(\gamma|\boldsymbol{y}, \boldsymbol{R}, \boldsymbol{\beta}, \tau)$$

which is sampled iteratively to form a dependent sample from the joint posterior distribution. The density $p(\tau|\boldsymbol{y}, \boldsymbol{R}, \boldsymbol{\beta}, \gamma)$ is an inverse gamma distribution, allowing τ to be integrated out of the full posterior in order to obtain the marginal posterior distribution $p(\boldsymbol{\beta}|\boldsymbol{y}, \boldsymbol{R}, \gamma)$. As all of the parameters have non-standard posterior densities, we use the Metropolis–Hastings (MH) algorithms (Hastings, 1970; Metropolis, Rosenbluth, Rosenbluth, & Teller, 1953).

In order to speed convergence and allow optimal mixing properties, we use the combined Random Walk and Independent Kernel MH algorithms. The Random Walk Metropolis algorithm is used for the first M iterations, the so-called burn-in period, while the Independent Kernel MH algorithm is employed from iteration $M + 1$ onwards, employing the sample mean and covariance matrix of the burn-in iterates for each parameter grouping. This procedure is discussed in detail by Chen and So (2006).

Bayesian forecasts of VaR can be constructed via the MCMC sampling scheme. For each MCMC iterate of parameter values $\boldsymbol{\beta}^{(j)}$ and $\gamma^{(j)}$, $j = 1, \ldots, M$, a 1-day α-level VaR estimator is obtained by plugging $\boldsymbol{\beta}^{(j)}$ and $\gamma^{(j)}$ into the formula for $f_{n+1}(\boldsymbol{\beta})$, obtaining $f_{n+1}(\boldsymbol{\beta})^{(j)}$. Under the TRV model, this is:

$$\begin{aligned} f_{n+1}(\boldsymbol{\beta})^{(j)} &= \left(\beta_1^{(j)} + \beta_2^{(j)} f_n(\boldsymbol{\beta}) + \beta_3^{(j)} R_n\right) I(R_n \le \gamma^{(j)}) \\ &\quad + \left(\beta_4^{(j)} + \beta_5^{(j)} f_n(\boldsymbol{\beta}) + \beta_6^{(j)} R_n\right) \\ &\quad \times (1 - I(R_n \le \gamma^{(j)})). \end{aligned}$$

These iterated values $f_{n+1}(\boldsymbol{\beta})^{(j)}$ are then simply averaged over the iterates $j = M + 1, \ldots, N$, to obtain a posterior mean estimate $\widehat{\text{VaR}}_{n+1}$ that is a forecast of VaR, where the parameters have been integrated out in the MCMC sampling scheme.

3.3. Parametric GARCH estimation

The parametric GARCH models, labelled Model A above, are estimated by MCMC methods here, following the method of Chen et al. (2003). First, the standard prior choices are made, so that the usual stationarity and positivity conditions are enforced, i.e.:

$$\begin{aligned}&\alpha_0 > 0; \qquad 0 < \alpha_1 + \beta_1 < 1;\\ &\alpha_1, \beta_1 \geq 0; \qquad |\phi_1| < 1,\end{aligned} \tag{11}$$

and the degrees of freedom for the Student-t errors are restricted to be above 4, ensuring that the first four moments are finite. This is achieved by placing a flat prior over the parameters constrained to the region B, which is equivalent to Eq. (11), and defining $\eta^* = 1/\eta$ and using a flat prior $\eta^* \sim$ Unif(0, 0.25). Thus, the prior becomes:

$$p(\phi_1, \phi_1, \alpha_0, \alpha_1, \beta_1) = I(B) \times I(\eta > 4).$$

The likelihood is defined by the choice of error distribution, combined with the GARCH volatility equation.

Multiplying the likelihood and the prior gives the posterior density function (up to a proportionality constant). The standard Gaussian random walk Metropolis method is employed for the first M MCMC iterations (M is the size of the burn-in sample) for each of the parameter groups: (i) (ϕ_1, ϕ_1); (ii) $(\alpha_0, \alpha_1, \beta_1)$; and (iii) η^*; in turn. After the burn-in period, the sample mean and sample variance–covariance of the iterates for $(\alpha_0, \alpha_1, \beta_1)$ are collected. These are then used as the proposal mean and covariance matrix in an independent kernel Metropolis–Hastings method, with a Gaussian proposal distribution. The overall method is thus adaptive, because it learns from the burn-in period. This has the added advantage of capturing the posterior correlations among the $\boldsymbol{\alpha}$ in the burn-in period for use in the sample period proposal, which should also increase the mixing rate of the sampling scheme, and, since the burn-in sample's mean (now the proposal mean) is probably not too close to the boundaries in Eq. (11), the sampler should perform better in that region for these parameters. For more details of this method, see Gerlach and Chen (2008) or Chen et al. (2003).

We note that this method will only work if the MCMC sample has converged on and sufficiently covered the posterior within the burn-in period. Convergence is thus monitored heavily using trace and ACF plots, while the tuning algorithm will also help to ensure sufficient coverage of the posterior by moderating the acceptance rate of the Metropolis method. MCMC results and convergence are examined extensively by starting the scheme from many different and varied starting values.

3.4. Forecast evaluation

In this section we discuss assessing the accuracy of VaR estimates and forecasts. The Basel II Accord requires financial institutions to use back-testing, so that at least one year of actual returns are compared with VaR forecasts. There are some common criteria for comparing the forecasting performance of VaR models; that is, the violations ($I(y_t < -\text{VaR}_t)$) and the violation rate (VRate). For an in-sample period of size n and forecast sample of size m, VRate is defined as:

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

which is simply the proportion of violating returns. Naturally, the VRate should be close to the risk level, α, for accurate risk models.

Three formal back-testing methods for assessing the forecasting performance are the unconditional coverage (uc) test of Kupiec (1995), the conditional coverage (cc) test of Christoffersen (1998), and the dynamic quantile (DQ) test of Engle and Manganelli (2004). Gaglianone, Lima, Linton, and Smith (2011) recently proposed another formal test, but our focus is on the three well-known versions just mentioned. Under the null hypothesis $\alpha = \alpha_0$, Kupiec (1995) employs a likelihood ratio to test whether VaR estimates, on average, provide a correct coverage of the lower α% tails of the forecast distributions. Christoffersen (1998) develops an independence test, employing a two-state Markov process, and combines this with the uc test to develop a joint likelihood ratio conditional coverage test that examines whether VaR estimates display correct conditional coverage at each point in time. Thus, the conditional coverage test simultaneously examines whether the violations appear independently and whether the unconditional coverage is α. The DQ test is also a joint test of the independence of violations and correct coverage. It employs a regression-based model of the violation-related variable 'hits', defined as $I\left(y_t < -\text{VaR}_t\right) - \alpha$, which will, on average, be α if the unconditional coverage is correct. A regression-type test is then employed to examine whether the 'hits' are related to lagged 'hits', lagged VaR forecasts, or other relevant regressors, over time; a model producing accurate and independent violations and 'hits' will not be. The DQ test is well known to be more powerful than the CC test, see e.g. Berkowitz, Christofferson, and Pelletier (2011).

The Basel II Accord stipulates that market risk charges (MRC) (also called Daily Capital Charges) should also be used to assess appropriate risk models where a lower MRC is desirable. The optimization problem facing ADIs, with the number of violations and forecasts of risk as endogenous choice variables, is as follows:

$$\text{Daily Capital Charge}_t = \sup\left\{\text{VaR}_{t-1}, (3+k)\overline{\text{VaR}_{60}}\right\},$$

where VaR_{t-1} is the VaR of the previous trading day, $\overline{\text{VaR}_{60}}$ is the average VaR over the last 60 trading days, and k is the penalty term from the Basel Accord Penalty Zone. The daily capital charge is set to be the supremum of the last trading day VaR and the average VaR over the past 60 trading days, multiplied by a violation penalty weight factor $(3 + k)$. Models with lower daily capital charge values are preferred for risk management. The daily capital charge attempts to give a conservative estimate of the capital required to cover market risk, and tries to correct for the under-estimation of risk levels by applying a penalty factor to the average of previous VaR estimates. The penalty is higher the more risk has been under-estimated in the past. MRC is the average of the daily capital charges during the forecast period.

McAleer and da Veiga (2008a, Table IV) display the penalty zones at the 1% level, where the number of violations is given for 250 trading days. We extend this table here to account for the large number of periods in our forecast sample. Our motivation for this is the following quote:

Table 1
Modified traffic light approach (Basel Committee, 1996) based on 400 trading days; the true coverage is 99%.

Zone	Number of violations	Cumulative probability	Increase in scaling factor
Green	0	0.01795	0
	1	0.09048	0
	2	0.23663	0
	3	0.43249	0
	4	0.62884	0
	5	0.78592	0
	6	0.89037	0
	7	0.94976	0
Yellow	8	0.97923	0.39820
	9	0.99220	0.48142
	10	0.99732	0.56080
	11	0.99915	0.63705
	12	0.99975	0.71069
Red	13 or more	0.99993	1

> "Another feature of regulatory back-tests that is not easy to understand is why they require only 250 days in the back-test. With such a small sample the power of the test to reject a false hypothesis is very low indeed. So, all in all, it is highly likely that an inaccurate VaR model will pass the regulatory backtest" (Alexander, 2008, p. 336).

We thus increase the usual forecast sample size here, and subsequently extend the traffic light approach to obtain the penalty weight factor for such larger samples, which are given in Table 1. This is discussed in more detail in Section 4.

Note that if models consistently under-estimate risk, and thus have too many violations, it is likely that they will have smaller values of MRC. On the other hand, models that consistently over-estimate risk levels for violations will have very large MRC values. As such, models with small MRC values are only preferred if they generate independent violations at the correct rate α.

Finally, the accuracy of quantile forecasts can be assessed directly using the quantile criterion loss function, given in Eq. (9). Here, $f_t(\boldsymbol{\beta})$ is replaced by the forecast VaRs for each method. The true VaR series should give the minimum of Eq. (9), and thus the most accurate model under consideration over the forecast period should return the minimum value of this loss function.

4. Empirical applications

In order to demonstrate and compare the forecasting performances of the proposed models, we first consider daily financial returns from five Asia-Pacific Economic Cooperation (APEC) stock markets: Standard and Poor's 500 Composite Index (US), Nikkei 225 Index (Japan), TAIEX Index (Taiwan), HANG SENG Index (Hong Kong) and KOSPI Index (Korea). An equally-weighted daily return portfolio is formed from the five individual market returns, on days when all markets traded. For this portfolio, the range data from Standard and Poor's 500 Composite Index is used in Eqs. (6)–(8). This choice is based on the global economic scale of the US market and its strong impact on the economic growth of other countries. Moreover, we also consider stock market returns and their intra-day ranges. All of the market data are obtained from Datastream International for the period January 1, 2002–April 30, 2010. A further example concerns two exchange rate series, the Euro vs. US and Japan vs. US exchange rates. The range data for the exchange rate series were obtained from the Thomson Reuters Tick History database.

The percentage returns series are calculated by taking differences of the logarithms or the daily price indices, $r_{t,j} = (\ln(P_t^j) - \ln(P_{t-1}^j)) \times 100$, where P_t^j is the closing price index of asset j on day t. We consider a single equal-weighted portfolio of assets, with returns: $y_t = \sum_{i=1}^{5} r_{t,i} \times 20\%$. In addition, the differences of the logarithm of the daily series of intra-day high and low prices are taken as the range data, R_t, and are defined as $R_t = (\ln(P_{t,\max}) - \ln(P_{t,\min})) \times 100$. We use the S&P 500 range for the portfolio as the explanatory variable, and the threshold variable, in the risk models using range data, and use the domestic daily range for this purpose in each individual stock market.

We now examine the performances of risk management strategies during the 2008–09 GFC, evaluate the way in which the financial crisis affects risk management practices, and forecast VaR and daily capital charges; that is, the diversification of more than a single investment for the purpose of risk control and management.

4.1. VaR forecasting for the portfolio

The full sample is divided into an in-sample period (from January 1, 2002 to July 31, 2008), and a forecast period of 400 trading days (from August 1, 2008 to April 30, 2010), which covers the 2008–09 GFC.

Fig. 1 shows the time series plots of the portfolio returns and the S&P500 range data; both highlight sharp increases in volatility in September 2008 and for the subsequent early months of 2009, plus a very low volatility period from approximately 2005–2007; the range data clearly reflect these periods well. Table 2 presents summary statistics for each market and the portfolio returns for both the full sample (including both in- and out-of-sample) and the forecast sample (out-of-sample). As expected, the forecast period displays consistently higher standard deviations and average intra-day ranges than the full sample, across

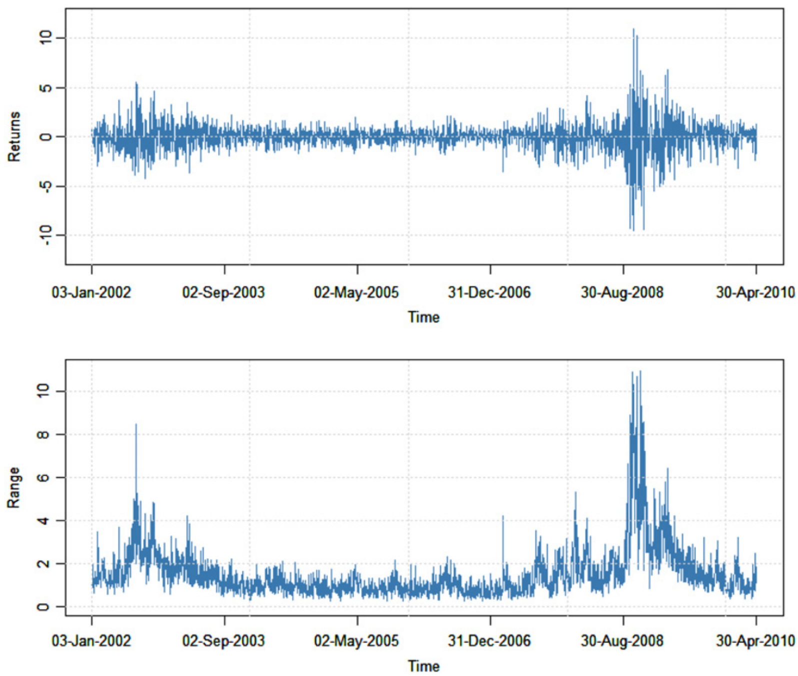

Fig. 1. Time series plots of (a) the portfolio's returns and (b) the S&P500 range.

Table 2
Summary statistics: stock index returns and ranges for five stock markets and an equal weights portfolio from January 1, 2002 to April 30, 2010.

Statistics	TAIEX		Nikkei225		HSI		KOSPI		S&P500		Portfolio	
	Return	Range	Return	Range	Return	Range	Return	Range	Return	Range	Return	Range
Observations	2055	2055	2042	2042	2056	2056	2062	2062	2096	2096	1790	1790
Mean	0.017	1.494	0.001	1.506	0.030	1.467	0.043	1.770	0.001	1.499	0.005	1.504
Median	0.064	1.277	0.044	1.287	0.055	1.188	0.150	1.531	0.074	1.167	1.164	1.167
Std.	1.478	0.852	1.618	0.988	1.662	1.072	1.639	1.097	1.391	1.185	0.056	1.174
Minimum	−6.912	0.146	−12.111	0.299	−13.582	0.285	−11.172	0.408	−9.470	0.239	−7.336	0.239
Maximum	6.525	7.403	13.235	11.743	13.407	17.647	11.284	15.841	10.957	10.904	8.297	10.904
$Q1$	−0.664	0.885	−0.781	0.905	−0.657	0.846	−0.726	1.092	−0.587	0.786	−0.539	0.797
$Q3$	0.793	1.851	0.875	1.844	0.785	1.761	0.931	2.108	0.606	1.811	0.617	1.818
Skewness	−0.257	1.653	−0.377	3.527	0.100	4.378	−0.453	3.715	−0.144	3.162	−0.297	3.009
Normality test	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Excess kurtosis	2.316	4.264	7.433	23.035	9.345	39.355	4.581	28.018	9.155	15.320	6.531	14.031
Hold-out set												
Mean	0.023	1.802	−0.050	1.914	−0.003	2.252	0.032	2.109	−0.013	2.332	−0.035	2.306
Std.	1.825	1.035	2.360	1.558	2.578	1.631	2.066	1.690	2.241	1.849	1.708	1.760
Minimum	−5.933	0.333	−12.111	0.314	−13.582	0.462	−11.172	0.408	−9.470	0.375	−7.336	0.375

all markets. Also as expected, all six return series have heavy-tailed distributions, and most are mildly negatively skewed. The p-values of the Jarque–Bera test for departure from normality are all very small: normality is rejected in all markets/series.

The specific VaR models and methods considered are:

1. Non-parametric: short-term (ST, last 25 days) and long-term (LT, last 100 days) sample percentiles.
2. Parametric methods: GARCH(1, 1) with normal and Student-t errors; RiskMetrics with normal errors and $\lambda = 0.94$.
3. Semi-parametric methods: The family of CAViaR models in Eqs. (1)–(8) is considered, with estimation by either frequentist (denoted "E&M", to indicate that an estimation the same as or similar to that of Engle & Manganelli, 2004, was used) and/or Bayesian methods, both as detailed in Section 3.

For each model and method considered, we generate one-day-ahead forecasts of VaR for each day in the forecast sample period. A rolling window approach is used, where a fixed in-sample size of approximately 2000 days is employed for estimation, in order to forecast each day in the forecast period. Thus, each method and/or model is completely re-estimated for each day in the forecast sample, employing the previous 2000 days as an input to each estimation. This gives each method a chance to adapt to the changing risk dynamics and levels.

We use Fortran codes to obtain the MCMC iterates, and use the '`fminsearch`' routine in the Matlab software to minimise Eq. (9) numerically. The Matlab code is adapted and updated for the RV-E&M model using freely available

Table 3
Bayesian estimation of parameters for the RV, TRV and TRIG specifications.

Model	Parameter	1%			5%		
		Mean	Std.	95% CI	Mean	Std.	95% CI
RV	β_1	0.233	0.042	(0.163, 0.333)	0.251	0.052	(0.157, 0.360)
	β_2	0.608	0.035	(0.521, 0.672)	0.339	0.055	(0.235, 0.445)
	β_3	0.548	0.044	(0.457, 0.647)	0.582	0.047	(0.490, 0.671)
TRV	β_1	0.221	0.036	(0.136, 0.284)	0.125	0.069	(−0.003, 0.270)
	β_2	0.850	0.029	(0.788, 0.905)	0.546	0.104	(0.360, 0.775)
	β_3	−0.030	0.057	(−0.132, 0.092)	0.440	0.125	(0.168, 0.659)
	β_4	1.541	0.196	(1.096, 1.897)	0.497	0.182	(0.139, 0.846)
	β_5	0.237	0.073	(0.070, 0.375)	0.162	0.090	(0.015, 0.360)
	β_6	0.400	0.051	(0.319, 0.517)	0.616	0.087	(0.426, 0.778)
	γ	1.447	0.004	(1.435, 1.452)	1.424	0.047	(1.339, 1.543)
TRIG	β_1	0.335	0.053	(0.237, 0.444)	0.170	0.082	(0.030, 0.351)
	β_2	0.802	0.025	(0.749, 0.847)	0.514	0.108	(0.287, 0.737)
	β_3	0.044	0.036	(0.001, 0.140)	0.544	0.167	(0.191, 0.838)
	β_4	4.604	0.518	(3.526, 5.553)	1.008	0.405	(0.282, 1.811)
	β_5	0.359	0.079	(0.180, 0.497)	0.162	0.094	(0.013, 0.371)
	β_6	0.405	0.066	(0.312, 0.569)	0.614	0.091	(0.406, 0.784)
	γ	1.447	0.004	(1.434, 1.452)	1.429	0.039	(1.365, 1.548)

code which was kindly provided by Simone Manganelli (downloadable from http://www.simonemanganelli.org/Simone/Research.html). The Econometrics toolbox in the software Matlab is employed for estimating both GARCH models via maximum likelihood.

For the Bayesian estimation, the priors are as stated in Section 3, e.g., a uniform prior is used for the threshold value γ; that is, $\gamma \sim \text{Unif}(l, u)$, where l and u are the 1st and 3rd quantiles of the range data. MCMC sampling is performed with a total of 20,000 iterations, including the first 10,000 burn-in iterations. The last 10,000 iterations are used for inference.

We only report Bayesian estimates of the parameters for the RV, TRV and TRIG specifications in Table 3, including posterior means, standard deviations (Std.), and the 95% credible interval (95% CI) for each parameter. All of the estimated parameters are significant, except for β_3 of the TRV model at the 1% level and β_1 of the TRV model at the 5% level. The former belongs to the low range which does not respond strongly to VaR, and the latter is the intercept of the low range. Convergence of the MCMC iterates is examined via trace plots and autocorrelation function plots as diagnostic checks, but these are not presented here due to space limitations; they show that the Markov chain appears to have reached a stationary distribution in each case, and indicate low autocorrelation and fairly efficient sampling; hence, they suggest good convergence and mixing properties of the MCMC sampling scheme.

The traffic light approach suggested by the Basel Committee (1996) deems a VaR model acceptable (green zone) if the number of violations of the 1% VaR remains below the binomial ($p = 0.01$) 95% quantile. A model is disputable (yellow zone) up to the 99.99% quantile, and is deemed seriously flawed (red zone) whenever more violations occur. Translated to our sample size ($n = 400$) in Table 1, a model passes a regulatory performance assessment if no more than seven violations occur, is disputable when between eight and 12 violations occur, and is seriously flawed for more than 12 violations.

The results reported in Table 4 show numbers of violations, zone colour, VRate, VRate/α, penalty charge, MRC and the quantile criterion function, at the 1% and 5% confidence levels for the portfolio return series, and for each of the 17 VaR models/methods considered; these are informal comparison metrics. When comparing VRate/α results, a value of 0.9 is considered better than VRate/$\alpha = 1.1$, as the loss estimates are conservative in the former case and anti-conservative in the latter. The methods are formally tested by UC, CC and DQ, with the results given in Table 7 (bottom right corner). At the 1% confidence level, the ST, LT, RM, both GARCHs, both SAVs, the AS-E&M and the TIG-Bayesian methods are rejected by at least one test, at the 5% significance level. None of the models using range information can be rejected, while all the of models which survive the tests are CAViaR-type models. In Table 4, the three top ranked models for each metric, of those surviving the tests, are boxed. Of these, the RV-Bayesian, TRV-Bayesian and TRIG-Bayesian are the top three according to VRate, while these plus RV-E&M are all in the green zone and attract no penalty. The RV, TRV, and TRIG models with the Bayesian approach have better performances based on the penalty, daily capital charge and loss function criteria. Informally, these are the most accurate VaR forecast methods for the portfolio 1% risk level.

At the 5% confidence level, the ST, LT, RM, both GARCHs and both SAV models are again rejected by at least one test, as are all of the CAViaR models estimated via the E&M method. As such, only the models estimated via Bayesian methods cannot be rejected in any of the tests conducted here. For these models, the TIG, TRIG, RV and TRV-Bayesian models are closest to nominal in terms of VRate/α. Across the metrics, the RV-Bayesian and TRV-Bayesian also rank best for the loss function values. Note that the Basel Accord only has 1% level penalties for MRC, not 5%, and thus these metrics are not reported in this case.

Table 5 shows the VaR forecasting results separately for each of the five individual markets making up the portfolio. For the Nikkei225, at the 1% risk level, only three of the methods are not rejected by any of the tests, as shown in Table 7: IG-Bayesian, TIG-Bayesian and RV-Bayesian. All of the models under-predict risk for the

Table 4
VaR prediction performance: using 17 model specifications and the 400 forecasts for the portfolio returns.

$\alpha = 1\%$	Violations	Zone	VRate/α	Penalty	Daily capital charge	Quantile criterion function
ST	23	Red	5.75	1.000	12.507	28.360
LT	11	Yellow	2.75	0.631	15.069	27.312
RiskMetrics	13	Red	3.25	0.762	14.884	20.837
GARCH-n	11	Yellow	2.75	0.631	13.014	20.430
GARCH-t	8	Yellow	2.00	0.400	13.113	19.329
SAV-Bayesian	8	Yellow	2.00	0.400	14.307	20.810
SAV-E&M	13	Red	3.25	0.762	13.278	23.462
AS-Bayesian	7	Green	1.75	0.150	$\boxed{10.900}$	23.636
AS-E&M	13	Red	3.25	0.762	12.428	25.786
IG-Bayesian	7	Green	1.75	0.150	12.283	18.901
IG-E&M	7	Green	1.75	0.150	12.229	18.684
TCAV-Bayesian	7	Green	1.75	0.150	$\boxed{11.104}$	23.022
TIG-Bayesian	8	Yellow	2.00	0.400	12.723	20.177
RV-Bayesian	3	Green	$\boxed{0.75}$	$\boxed{0.000}$	12.124	$\boxed{17.388}$
RV-E&M	6	Green	1.50	$\boxed{0.000}$	$\boxed{11.049}$	$\boxed{17.092}$
TRV-Bayesian	3	Green	$\boxed{0.75}$	$\boxed{0.000}$	11.719	$\boxed{17.152}$
TRIG-Bayesian	4	Green	$\boxed{1.00}$	$\boxed{0.000}$	11.545	18.459
$\alpha = 5\%$	Violations	VRate	VRate/α			Quantile criterion function
ST	38	9.50	1.90			28.360
LT	21	5.25	1.05			27.312
RiskMetrics	26	6.50	1.30			20.837
GARCH-n	29	7.25	1.45			20.430
GARCH-t	30	7.50	1.50			19.329
SAV-Bayesian	26	6.50	1.30			20.810
SAV-E&M	30	7.50	1.50			23.462
AS-Bayesian	25	6.25	1.25			23.636
AS-E&M	28	7.00	1.40			25.786
IG-Bayesian	24	6.00	1.20			18.901
IG-E&M	25	6.25	1.25			18.684
TCAV-Bayesian	23	5.75	1.15			23.022
TIG-Bayesian	20	5.00	$\boxed{1.00}$			20.177
RV-Bayesian	22	5.50	$\boxed{1.10}$			$\boxed{17.388}$
RV-E&M	23	5.75	1.15			17.092
TRV-Bayesian	22	5.50	$\boxed{1.10}$			$\boxed{17.152}$
TRIG-Bayesian	21	5.25	$\boxed{1.05}$			$\boxed{18.459}$

Nikkei, since all of the VRates are larger than 0.01. The IG-Bayesian and RV-Bayesian do best (of all models) on VRate, but of these three models, IG-Bayesian does best on four of the six criteria (including 'Zone'). For the HSI index returns, all of the models under-predict risk levels, but only the ad hoc ST and LT methods fail the tests. Among the remaining models, the AS-Bayesian, TCAV-Bayesian and SAV-Bayesian consistently rank in the top three across all criteria, for HSI. For the KOSPI, only the SAV, AS, IG, RV and TRV, all estimated via the Bayesian method, are not rejected by the tests. Of these, the SAV, TIG and RV-Bayesian all consistently rank in the top three over all metrics. For the TAIEX, only the SAV, IG and RV models, using both E&M and Bayesian estimation, are not rejected by the tests, and only the RV-Bayesian is a conservative risk model (with VRate < 0.01). Of these, the IG-Bayesian and IG-E&M consistently rank well across the criteria. Finally, for the S&P500, no model survives the tests at the 1% VaR level.

A similar story appears at the 5% level; see Table 6. All of the models are rejected for the Nikkei225, though not at the 1% significance level for TCAV-Bayesian and TIG-Bayesian, which also perform comparatively well across the criteria in Table 5. For the HSI, the ST, LT, GARCH-t, TIG-Bayesian and TRIG-Bayesian are all rejected. For the other models, the AS-Bayesian, TCAV-Bayesian and TRV-Bayesian rank best across the criteria. For the KOSPI and the TAIEX, the IG-Bayesian and RV-Bayesian ranked most consistently among the models which were not rejected by the tests, while the TIG-Bayesian ranked best among the surviving models for the S&P500.

From above, we note that the CAViaR family of models consistently dominate the models that are not rejected the formal back-tests, while traditional methods such as RM and GARCH are consistently rejected. Further, CAViaR models estimated by the Bayesian method consistently rank higher than models estimated by the traditional method in E&M, across all or most series. Finally, models with the IG or RV form consistently survive the tests and rank highly for most series.

We consider two exchange rates, namely the Euro vs. US and Japan vs. US exchange rates. Due to the availability of intra-day data for these series, the dates for the sample and forecast periods are December 21, 2004–July 3, 2009, and July 6, 2009–February 8, 2011. While it is well known that exchange rate returns do not generally exhibit significant

Table 5
VaR prediction performance over 450 forecasts at the 1% level for each market.

$\alpha = 1\%$		Violations	Zone	VRate/α	Daily capital	Quantile criterion function
Nikkei225	ST	20	Red	4.44	16.252	35.920
	LT	11	Yellow	2.44	19.008	40.879
	RiskMetrics	8	Yellow	1.78	15.060	31.756
	GARCH-t	5	Green	1.11	14.556	29.227
	GARCH-n	6	Green	1.33	15.407	29.657
	SAV-Bayesian	8	Yellow	1.78	15.776	28.659
	SAV-E&M	14	Red	3.11	18.139	29.605
	AS-Bayesian	9	Yellow	2.00	17.352	28.867
	AS-E&M	14	Red	3.11	17.761	32.905
	IG-Bayesian	7	Green	1.56	15.074	27.009
	IG-E&M	9	Yellow	2.00	16.605	26.860
	TCAV-Bayesian	10	Yellow	2.22	17.288	30.144
	TIG-Bayesian	8	Yellow	1.78	14.562	29.212
	RV-Bayesian	7	Green	1.56	14.864	27.455
	RV-E&M	8	Yellow	1.78	14.218	27.305
	TRV-Bayesian	9	Yellow	2.00	17.317	27.279
	TRIG-Bayesian	9	Yellow	2.00	16.685	29.158
HSI	ST	21	Red	4.67	17.194	39.719
	LT	10	Yellow	2.22	18.255	39.872
	RiskMetrics	5	Green	1.11	16.995	30.880
	GARCH-t	4	Green	0.89	16.449	30.401
	GARCH-n	7	Green	1.56	17.952	30.263
	SAV-Bayesian	5	Green	1.11	17.187	26.822
	SAV-E&M	4	Green	0.89	17.303	27.743
	AS-Bayesian	3	Green	0.67	15.986	24.628
	AS-E&M	2	Green	0.44	16.587	25.235
	IG-Bayesian	5	Green	1.11	16.848	26.750
	IG-E&M	6	Green	1.33	16.035	27.337
	TCAV-Bayesian	6	Green	1.33	15.454	25.140
	TIG-Bayesian	6	Green	1.33	15.337	25.916
	RV-Bayesian	3	Green	0.67	17.984	28.764
	RV-E&M	5	Green	1.11	15.990	29.134
	TRV-Bayesian	5	Green	1.11	17.570	28.365
	TRIG-Bayesian	3	Green	0.67	17.820	28.507
KOSPI	ST	22	Red	4.89	15.966	39.185
	LT	9	Yellow	2.00	17.816	39.794
	RiskMetrics	12	Yellow	2.67	16.081	33.103
	GARCH-t	10	Yellow	2.22	15.402	32.545
	GARCH-n	11	Yellow	2.44	16.236	30.831
	SAV-Bayesian	8	Yellow	1.78	14.933	31.693
	SAV-E&M	10	Yellow	2.22	16.029	30.051
	AS-Bayesian	8	Yellow	1.78	14.326	31.417
	AS-E&M	13	Yellow	2.89	15.096	33.246
	IG-Bayesian	8	Yellow	1.78	15.062	30.435
	IG-E&M	10	Yellow	2.22	16.543	30.217
	TCAV-Bayesian	10	Yellow	2.22	16.485	31.879
	TIG-Bayesian	10	Yellow	2.22	16.704	32.655
	RV-Bayesian	8	Yellow	1.78	15.244	31.346
	RV-E&M	10	Yellow	2.22	15.849	31.328
	TRV-Bayesian	8	Yellow	1.78	15.541	32.951
	TRIG-Bayesian	6	Green	1.33	16.276	33.138
TAIEX	ST	21	Red	4.67	13.911	32.464
	LT	11	Yellow	2.44	14.931	26.866
	RiskMetrics	11	Yellow	2.44	14.499	25.582
	GARCH-t	8	Yellow	1.78	14.285	25.145
	GARCH-n	11	Yellow	2.44	13.439	23.812
	SAV-Bayesian	6	Green	1.33	13.415	22.923
	SAV-E&M	7	Green	1.56	13.174	22.844
	AS-Bayesian	9	Yellow	2.00	14.649	24.434
	AS-E&M	13	Yellow	2.89	14.243	27.034
	IG-Bayesian	6	Green	1.33	13.728	23.242
	IG-E&M	6	Green	1.33	13.563	23.114
	TCAV-Bayesian	11	Yellow	2.44	15.369	23.597
	TIG-Bayesian	9	Yellow	2.00	15.043	25.371
	RV-Bayesian	4	Green	0.89	13.690	23.308
	RV-E&M	5	Green	1.11	13.487	23.291
	TRV-Bayesian	6	Green	1.33	13.731	26.026
	TRIG-Bayesian	6	Green	1.33	13.695	25.123

Table 5 (*continued*)

$\alpha = 1\%$		Violations	Zone	VRate/α	Daily capital	Quantile criterion function
S&P500	ST	24	Red	5.33	15.903	39.708
	LT	11	Yellow	2.44	18.777	35.553
	RiskMetrics	13	Yellow	2.89	17.522	28.498
	GARCH-*t*	9	Yellow	2.00	18.200	28.410
	GARCH-*n*	14	Red	3.11	16.726	27.212
	SAV-Bayesian	10	Yellow	2.22	16.379	28.492
	SAV-E&M	13	Yellow	2.89	16.597	34.410
	AS-Bayesian	11	Yellow	2.44	16.450	28.306
	AS-E&M	15	Red	3.33	16.658	32.543
	IG-Bayesian	9	Yellow	2.00	16.369	28.084
	IG-E&M	11	Yellow	2.44	17.221	30.176
	TCAV-Bayesian	12	Yellow	2.67	17.123	31.511
	TIG-Bayesian	11	Yellow	2.44	17.029	28.016
	RV-Bayesian	9	Yellow	2.00	15.385	26.847
	RV-E&M	17	Red	3.77	16.061	31.357
	TRV-Bayesian	10	Yellow	2.22	15.409	29.033
	TRIG-Bayesian	11	Yellow	2.44	15.657	29.835

volatility asymmetry, we choose to keep the same set of 17 models/methods, including the asymmetric ones, for consistency. Table 8 contains the *p*-values for the tests for the 1% and 5% VaR forecasts for these two series over the 17 methods. Tables 9–10 show the violation rates and other accuracy metrics for these two series. At the 1% confidence level for the Euro/US series, only the two ad hoc methods, ST and LT, fail the tests. Of the surviving models, the AS-Bayesian, TCAV-E&M and TRIG-Bayesian consistently rank well across the metrics. For the JP/US rates at the 1% risk level, the ST, SAV-E&M, TCAV-E&M and RV-E&M models fail the tests. Of the surviving models, the TIG and TRIG, both Bayesian, and the GARCH-*t* model consistently rank in the top 3 across the various metrics.

At the 5% risk level, the ST, LT, RM, both GARCHs, both SAVs and both AS models are rejected, for the Euro/US series. Of the remaining models, the TIG and RV, both Bayesian, consistently rank highly. For the JP/US series, only the AS models, plus RV-Bayesian and TRV-Bayesian, survive the tests. Of these, AS-Bayesian and RV-Bayesian rank the highest across the metrics.

To summarise, similar conclusions apply for the exchange rate series: the CAViaR family of models are consistently the only models surviving the formal back-tests; CAViaR models estimated by the Bayesian method rank the highest across most series, and the models with the IG and RV forms consistently survive the tests and rank highly in most series.

Table 11 shows counts of rejections over the three tests (UC, CC, and DQ) and across the markets and exchange rate series for each model (counting the number of series where at least one test rejected each model). The DQ statistic is clearly the most powerful test and rejects the most models in the most markets. For the 1% VaR forecasts, the non-parametric ST and LT methods are rejected in all markets by almost all tests; they are clearly the poorest methods for this data period. The RM, GARCH with Gaussian and Student-*t* errors, and SAV-E&M and AS-E&M methods are rejected in five of the eight series. On the other hand, the RV-Bayesian and IG-Bayesian models are each rejected in only one series, the S&P500, and only by the powerful DQ test, while all other methods are rejected at least three times. The detection of violations in the portfolio return series using the TRIG-Bayesian model at the 1% level is shown in Fig. 2. There are four violations within the forecasting period, three of which occur during the period September–December 2008. The closeness of these violations in time is sufficient for the DQ test to reject this model. Other models met a similar fate due to not reacting enough, in time or magnitude, to the onset of the extreme GFC period. On the other hand, the RV-Bayesian and IG-Bayesian models were clearly able to do so effectively across almost all eight of these series.

For the 5% VAR forecasts, Table 11 shows that the ST, LT, GARCH-*t* and AS-E&M methods again perform the worst, being rejected in at least six series. Again, the IG-Bayesian and RV-Bayesian models perform the best, being rejected in only two of the eight series.

Table 12 shows summary statistics for the various criteria in Tables 4 and 5 for 1% and 5% VaR forecasting. While these are informal criteria, this table highlights the performances of each model across the six series, and may show consistent out-performance or the reverse. For each criterion, the means and medians across the eight series are shown. For the VRate/α, the square root of the average squared distance from 1 is also shown, and is labeled 'RMSD'. The best three models are boxed for each criterion summary, while the worst model appears in bold. For 1% VaR forecasting, the non-parametric models are clearly anti-conservative, under-estimating the risk levels and performing the worst among these 17 methods across the six series. The CAViaR models clearly perform the best as a group, since they get all of the boxes except one, with two stand-outs across the criteria: the RV-Bayesian model, which is in the top three ranked models for six of the seven criterion summaries; and the IG-Bayesian model, which is in the top three ranked models for four of the criterion summaries.

The lower panel of Table 12 shows these summaries for 5% VaR forecasting. Again, the RV-Bayesian, with four rankings in the top three across the five criteria, consistently does best across the criterion summaries over the eight series combined.

Table 6
VaR prediction performance over 450 forecasts at the 5% level for each market.

$\alpha = 5\%$		Violations	VRate	VRate/α	Quantile criterion function
Nikkei225	ST	39	8.67	1.73	35.920
	LT	20	4.44	0.89	40.879
	RiskMetrics	34	7.56	1.51	31.756
	GARCH-t	36	8.00	1.60	29.227
	GARCH-n	34	7.56	1.51	29.657
	SAV-Bayesian	29	6.45	1.29	28.659
	SAV-E&M	34	7.56	1.51	29.605
	AS-Bayesian	32	7.11	1.42	28.867
	AS-E&M	33	7.33	1.47	32.905
	IG-Bayesian	32	7.11	1.42	27.009
	IG-E&M	32	7.11	1.42	26.860
	TCAV-Bayesian	28	6.22	1.24	30.144
	TIG-Bayesian	28	6.22	1.24	29.212
	RV-Bayesian	33	7.33	1.47	27.455
	RV-E&M	36	8.00	1.60	27.305
	TRV-Bayesian	30	6.67	1.33	27.279
	TRIG-Bayesian	34	7.56	1.51	29.158
HSI	ST	40	8.89	1.78	39.719
	LT	25	5.56	1.11	39.872
	RiskMetrics	22	4.89	0.98	30.880
	GARCH-t	27	6.00	1.20	30.401
	GARCH-n	25	5.56	1.11	30.263
	SAV-Bayesian	24	5.33	1.04	26.822
	SAV-E&M	22	4.89	0.98	27.743
	AS-Bayesian	28	6.22	1.24	24.628
	AS-E&M	22	4.89	0.98	25.235
	IG-Bayesian	25	5.55	1.11	26.750
	IG-E&M	24	5.33	1.07	27.337
	TCAV-Bayesian	27	6.00	1.20	25.140
	TIG-Bayesian	29	6.44	1.29	25.916
	RV-Bayesian	22	4.89	0.98	28.764
	RV-E&M	22	4.89	0.98	29.134
	TRV-Bayesian	25	5.56	1.11	28.365
	TRIG-Bayesian	28	6.22	1.24	28.507
KOSPI	ST	38	8.44	1.69	39.185
	LT	22	4.88	0.98	39.794
	RiskMetrics	30	6.67	1.33	33.103
	GARCH-t	30	6.67	1.33	32.545
	GARCH-n	28	6.22	1.24	30.831
	SAV-Bayesian	28	6.22	1.24	31.693
	SAV-E&M	30	6.67	1.33	30.051
	AS-Bayesian	28	6.22	1.24	31.417
	AS-E&M	28	6.22	1.24	33.246
	IG-Bayesian	27	6.00	1.20	30.435
	IG-E&M	26	5.78	1.16	30.217
	TCAV-Bayesian	28	6.22	1.24	31.879
	TIG-Bayesian	26	5.78	1.16	32.655
	RV-Bayesian	23	5.11	1.02	31.346
	RV-E&M	23	5.11	1.02	31.328
	TRV-Bayesian	23	5.11	1.02	32.951
	TRIG-Bayesian	24	5.33	1.07	33.138
TAIEX	ST	44	9.78	1.96	32.464
	LT	25	5.56	1.11	26.866
	RiskMetrics	29	6.44	1.29	25.582
	GARCH-t	29	6.44	1.29	25.145
	GARCH-n	28	6.22	1.24	23.812
	SAV-Bayesian	26	5.78	1.56	22.923
	SAV-E&M	28	6.22	1.24	22.844
	AS-Bayesian	29	6.44	1.29	24.434
	AS-E&M	33	7.33	1.47	27.034
	IG-Bayesian	25	5.56	1.11	23.242
	IG-E&M	27	6.00	1.20	23.114
	TCAV-Bayesian	30	6.67	1.33	23.597
	TIG-Bayesian	31	6.89	1.38	25.371
	RV-Bayesian	23	5.11	1.02	23.308
	RV-E&M	26	5.78	1.16	23.291
	TRV-Bayesian	24	5.33	1.07	26.026
	TRIG-Bayesian	26	5.78	1.56	25.123

Table 6 (*continued*)

$\alpha = 5\%$		Violations	VRate	VRate/α	Quantile criterion function
S&P500	ST	43	9.56	1.91	39.708
	LT	29	6.44	1.29	35.553
	RiskMetrics	29	6.44	1.29	28.498
	GARCH-*t*	31	6.89	1.38	28.410
	GARCH-*n*	31	6.89	1.38	27.212
	SAV-Bayesian	31	6.89	1.38	28.492
	SAV-E&M	30	6.67	1.33	34.410
	AS-Bayesian	31	6.89	1.38	28.306
	AS-E&M	40	8.89	1.78	32.543
	IG-Bayesian	29	6.44	1.29	28.084
	IG-E&M	31	6.89	1.38	30.176
	TCAV-Bayesian	30	6.67	1.33	31.511
	TIG-Bayesian	30	6.67	1.33	28.016
	RV-Bayesian	33	7.33	1.47	26.847
	RV-E&M	41	9.11	1.82	31.357
	TRV-Bayesian	34	7.56	1.51	29.033
	TRIG-Bayesian	33	7.33	1.51	29.835

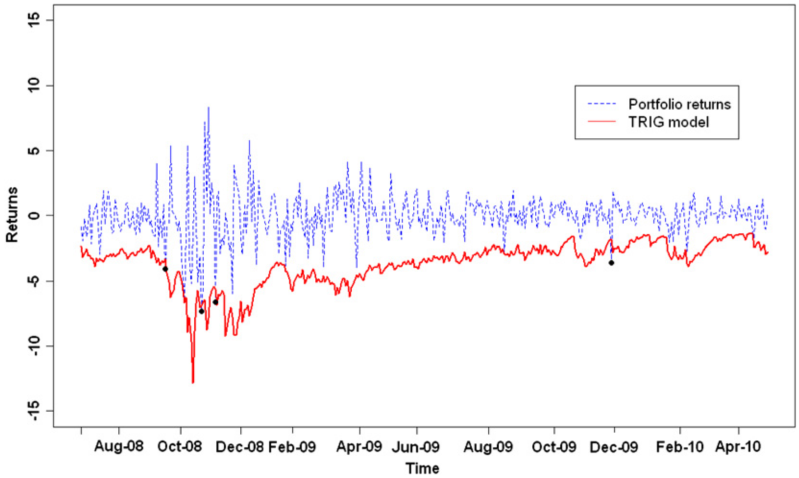

Fig. 2. VaR forecasts of the TRIG-Bayesian model at the 1% level.

4.2. Summary and discussion

Overall, the non-parametric methods were consistently the worst at VaR forecasting in each of the eight series, over a range of metrics. The ST method's performance can perhaps be explained by the fact that 25 days is not a sufficiently long period for estimating quantiles at the 1% and 5% risk levels; however, the LT method's poor performance is harder to explain. Clearly, these methods are not suitable for VaR forecasting at the 1% and 5% levels. Further, the fully parametric RiskMetrics and GARCH models only performed marginally better than the non-parametric ones, being rejected by at least one test in most or all series, and generally and consistently under-estimating risk levels and capturing risk dynamics poorly across the data analysed.

As a group, the CAViaR models performed uniformly better than the above-mentioned methods on all metrics for almost all series, and also for almost all metrics combined across the series, for both 1% and 5% VaR forecasting. Further, when focusing on models estimated by Bayesian or classical methods (i.e., E&M), the Bayesian models almost uniformly performed better than the same model estimated via E&M across all markets and metrics. Gerlach et al. (2011) found in simulations that the Bayesian estimates of the TCAV model parameters and forecasts of VaR were more efficient, i.e. had lower sampling variances, than those estimated/forecasted using classical estimation. The results here suggest that this is a more general result across the CAViaR family of models: Bayesian estimation and forecasting of CAViaR models is more accurate than classical estimation of CAViaR models, at least for the models and data considered here. Finally, three models stood out as performing the best across the back-tests and the range of forecast accuracy criteria applied: the IG-Bayesian and RV-Bayesian models for 1% VaR forecasting; and the TRV-Bayesian and RV-Bayesian models for 5% VaR forecasting. These models consistently out-performed all others across the eight series on most of the metrics considered, at each risk level.

The Basel Committee (1996) classified the reasons for model back-testing failures into the following categories:

1. Basic integrity of the model: The system is unable to capture the risk of the positions or there is a problem in calculating volatilities and correlations.
2. The model's accuracy could be improved: Risk of some instruments not measured with sufficient precision.

Table 7
p-values of the unconditional and conditional coverage tests, and dynamic quantile tests.

	Nikkei225						HSI					
	1%			5%			1%			5%		
	LR_{uc}	LR_{cc}	DQ	LR_{uc}	LR_{cc}	DQ	LR_{uc}	LR_{cc}	DQ	LR_{uc}	LR_{cc}	DQ
ST	0.000	0.000	0.000	0.001	0.003	0.000	0.000	0.000	0.000	0.001	0.002	0.000
LT	0.009	0.025	0.000	0.582	0.146	0.000	0.025	0.004	0.000	0.595	0.126	0.000
RiskMetrics	0.135	0.281	0.000	0.020	0.061	0.005	0.816	0.919	0.740	0.914	0.697	0.156
GARCH-n	0.499	0.731	0.000	0.020	0.032	0.004	0.273	0.489	0.331	0.595	0.756	0.194
GARCH-t	0.816	0.919	0.000	0.007	0.021	0.001	0.809	0.938	0.562	0.345	0.604	0.044
SAV-Bayesian	0.135	0.281	0.001	0.177	0.302	0.004	0.816	0.919	0.607	0.748	0.779	0.632
SAV-E&M	0.000	0.001	0.000	0.020	0.032	0.001	0.809	0.938	0.615	0.914	0.697	0.181
AS-Bayesian	0.061	0.142	0.043	0.053	0.090	0.006	0.449	0.739	0.564	0.251	0.414	0.073
AS-E&M	0.000	0.001	0.000	0.033	0.055	0.004	0.183	0.411	0.366	0.914	0.697	0.184
IG-Bayesian	0.273	0.489	0.243	0.053	0.090	0.006	0.816	0.919	0.653	0.595	0.756	0.214
IG-E&M	0.061	0.142	0.043	0.053	0.013	0.002	0.499	0.731	0.464	0.748	0.779	0.224
TCAV-Bayesian	0.025	0.063	0.007	0.251	0.079	0.047	0.499	0.731	0.595	0.345	0.542	0.085
TIG-Bayesian	0.135	0.281	0.132	0.251	0.079	0.037	0.499	0.731	0.283	0.177	0.395	0.026
RV-Bayesian	0.273	0.489	0.239	0.033	0.096	0.006	0.449	0.739	0.753	0.914	0.992	0.180
RV-E&M	0.135	0.281	0.000	0.007	0.021	0.000	0.816	0.919	0.432	0.914	0.697	0.176
TRV-Bayesian	0.061	0.142	0.025	0.122	0.210	0.009	0.816	0.919	0.535	0.595	0.756	0.027
TRIG-Bayesian	0.061	0.142	0.036	0.020	0.061	0.008	0.449	0.739	0.610	0.251	0.144	0.010
	KOSPI						TAIEX					
ST	0.000	0.000	0.000	0.002	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000
LT	0.061	0.064	0.000	0.914	0.697	0.001	0.009	0.018	0.000	0.595	0.806	0.101
RiskMetrics	0.003	0.009	0.000	0.122	0.297	0.158	0.009	0.025	0.000	0.177	0.284	0.056
GARCH-n	0.009	0.025	0.002	0.251	0.414	0.104	0.009	0.025	0.000	0.251	0.331	0.071
GARCH-t	0.025	0.063	0.007	0.122	0.210	0.094	0.135	0.281	0.000	0.177	0.284	0.061
SAV-Bayesian	0.135	0.281	0.111	0.251	0.414	0.183	0.499	0.731	0.438	0.460	0.694	0.196
SAV-E&M	0.025	0.063	0.007	0.122	0.210	0.104	0.273	0.489	0.275	0.251	0.331	0.070
AS-Bayesian	0.135	0.281	0.125	0.251	0.414	0.002	0.061	0.142	0.026	0.177	0.302	0.018
AS-E&M	0.001	0.003	0.000	0.251	0.414	0.001	0.001	0.003	0.000	0.033	0.055	0.000
IG-Bayesian	0.135	0.281	0.127	0.345	0.542	0.145	0.499	0.731	0.251	0.595	0.756	0.232
IG-E&M	0.025	0.063	0.007	0.460	0.678	0.198	0.499	0.731	0.278	0.345	0.604	0.168
TCAV-Bayesian	0.025	0.063	0.010	0.251	0.414	0.001	0.009	0.025	0.001	0.122	0.210	0.017
TIG-Bayesian	0.025	0.063	0.007	0.460	0.152	0.019	0.061	0.142	0.013	0.081	0.140	0.006
RV-Bayesian	0.135	0.281	0.088	0.914	0.287	0.220	0.809	0.938	0.177	0.914	0.978	0.516
RV-E&M	0.025	0.063	0.007	0.914	0.978	0.441	0.816	0.919	0.193	0.460	0.694	0.514
TRV-Bayesian	0.135	0.281	0.027	0.914	0.978	0.472	0.499	0.731	0.026	0.748	0.779	0.198
TRIG-Bayesian	0.499	0.731	0.300	0.748	0.912	0.287	0.499	0.731	0.022	0.460	0.694	0.369
	S&P500						Portfolio					
ST	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000
LT	0.009	0.025	0.000	0.177	0.271	0.000	0.004	0.009	0.000	0.820	0.267	0.000
RiskMetrics	0.001	0.003	0.000	0.177	0.312	0.035	0.000	0.001	0.000	0.187	0.400	0.040
GARCH-n	0.000	0.001	0.000	0.081	0.146	0.028	0.004	0.009	0.000	0.052	0.148	0.015
GARCH-t	0.061	0.142	0.000	0.081	0.146	0.028	0.077	0.176	0.022	0.032	0.096	0.007
SAV-Bayesian	0.025	0.063	0.000	0.081	0.146	0.006	0.077	0.070	0.000	0.187	0.400	0.092
SAV-E&M	0.001	0.003	0.000	0.122	0.218	0.009	0.000	0.001	0.000	0.032	0.059	0.005
AS-Bayesian	0.009	0.025	0.000	0.081	0.146	0.018	0.173	0.346	0.084	0.269	0.502	0.098
AS-E&M	0.000	0.000	0.000	0.001	0.003	0.000	0.000	0.001	0.000	0.083	0.164	0.011
IG-Bayesian	0.061	0.142	0.000	0.177	0.302	0.068	0.173	0.346	0.105	0.373	0.596	0.071
IG-E&M	0.009	0.025	0.000	0.081	0.146	0.012	0.173	0.346	0.079	0.269	0.502	0.044
TCAV-Bayesian	0.003	0.009	0.000	0.122	0.210	0.108	0.173	0.346	0.111	0.501	0.667	0.223
TIG-Bayesian	0.009	0.025	0.000	0.122	0.210	0.052	0.077	0.176	0.049	1.000	0.645	0.232
RV-Bayesian	0.061	0.142	0.000	0.033	0.055	0.024	0.599	0.853	0.642	0.651	0.878	0.295
RV-E&M	0.000	0.000	0.000	0.000	0.001	0.000	0.349	0.586	0.425	0.501	0.318	0.015
TRV-Bayesian	0.025	0.063	0.000	0.020	0.061	0.004	0.599	0.853	0.772	0.651	0.878	0.293
TRIG-Bayesian	0.009	0.025	0.000	0.033	0.055	0.023	1.000	0.960	0.561	0.820	0.302	0.177

3. Bad luck, or markets moved in a fashion that could not be anticipated by the model. For instance, volatilities or correlations turned out to be significantly different to what was predicted.
4. Intra-day trading: There was a change in positions after the VaR estimates were computed.

Of the markets considered here, only the US series was not able to be modeled effectively by at least one of the CAViaR models. The US market result, where all of the models are rejected for VaR 1% forecasting, may fit under point 3 above. We surmise that this may be the case, since the CAViaR models as a group were unable to be rejected for all of the series, except the US. Thus, we suspect that the US market alone "moved in a fashion that could not be anticipated" by the models we tried. This points to the possibility of future research on developing a model that can capture the tail risk dynamics in the US market, especially during the extreme or crisis conditions that prevailed during our forecast sample, since all of the models and methods in this paper failed at that task.

Table 8
p-values of the unconditional and conditional coverage tests, and dynamic quantile tests for each exchange rate series.

		1%			5%		
		LR_{uc}	LR_{cc}	DQ	LR_{uc}	LR_{cc}	DQ
Euro vs. US	ST	0.000	0.000	0.000	0.083	0.002	0.000
	LT	1.000	0.960	0.024	0.651	0.301	0.004
	RiskMetrics	0.629	0.833	0.735	0.651	0.301	0.033
	GARCH-n	1.000	0.960	0.698	0.651	0.301	0.043
	GARCH-t	0.266	0.536	0.591	0.373	0.314	0.046
	SAV-Bayesian	1.000	0.960	0.575	0.817	0.168	0.011
	SAV-E&M	0.629	0.833	0.674	0.817	0.168	0.011
	AS-Bayesian	1.000	0.960	0.425	1.000	0.220	0.009
	AS-E&M	1.000	0.960	0.432	1.000	0.220	0.011
	IG-Bayesian	0.349	0.586	0.427	0.373	0.596	0.054
	TCAV-Bayesian	1.000	0.960	0.763	0.641	0.385	0.111
	TCAV-E&M	1.000	0.960	0.816	0.342	0.586	0.200
	TIG-Bayesian	0.349	0.586	0.193	0.651	0.704	0.133
	RV-Bayesian	0.173	0.346	0.191	0.651	0.704	0.247
	RV-E&M	0.266	0.536	0.588	0.231	0.425	0.199
	TRV-Bayesian	0.349	0.586	0.430	0.817	0.557	0.112
	TRIG-Bayesian	0.349	0.586	0.408	0.820	0.696	0.103
JP vs. US	ST	0.004	0.011	0.000	0.373	0.596	0.000
	LT	0.266	0.536	0.584	1.000	0.348	0.195
	RiskMetrics	0.349	0.586	0.362	0.501	0.667	0.014
	GARCH-n	1.000	0.960	0.724	0.817	0.557	0.020
	GARCH-t	0.599	0.853	0.559	0.817	0.557	0.028
	SAV-Bayesian	0.599	0.853	0.589	0.817	0.557	0.008
	SAV-E&M	0.629	0.833	0.001	0.817	0.557	0.009
	AS-Bayesian	0.599	0.853	0.510	0.641	0.385	0.052
	AS-E&M	1.000	0.960	0.191	0.480	0.369	0.056
	IG-Bayesian	0.599	0.853	0.573	0.817	0.557	0.018
	TCAV-Bayesian	0.599	0.853	0.428	0.817	0.971	0.010
	TCAV-E&M	1.000	0.960	0.009	0.501	0.753	0.010
	TIG-Bayesian	0.599	0.853	0.602	0.817	0.377	0.025
	RV-Bayesian	0.599	0.853	0.575	0.817	0.971	0.125
	RV-E&M	0.629	0.833	0.000	0.641	0.447	0.013
	TRV-Bayesian	0.599	0.853	0.594	0.641	0.882	0.062
	TRIG-Bayesian	0.599	0.853	0.591	1.000	0.645	0.036

5. Concluding remarks

A novel family of risk models, namely nonlinear threshold CAViaR models using intra-day price ranges, is proposed, and the selection of optimal risk models during the 2008–09 global financial crisis is assessed and discussed. Risk management strategies and performances are assessed during this period, and the performances of VaR models during the global financial crisis are evaluated and compared. Furthermore, both Bayesian and frequentist methods of estimation and forecasting are assessed and compared with real financial return data. Bayesian MCMC methods are adapted to the new family of CAViaR models, employing the link between the quantile criterion function and the Skewed–Laplace density. Five APEC stock market indices are considered, both individually and via an equally weighted portfolio, and the VaR is forecast for these series over roughly a two year period. Two exchange rate series are also analysed in this manner. The empirical evidence reveals that risk levels and dynamics during the financial crisis seem to be predictable at a one day horizon, at least by some models (but not by others), in most markets, in the portfolio considered, and in the two exchange rate series. Also, by comparing the same model using different estimation methods, the forecasting performance of the Bayesian method is more accurate than that of the frequentist method, in almost all cases. Further, the forecasting performance of VaR models with range information outperforms that of models without range information, in most cases. Semi-parametric models, i.e. CAViaR, consistently ranked best via statistical testing and informal assessment criteria for VaR forecasting; next came the parametric models, though they were consistently rejected by back-tests, while the non-parametric methods consistently ranked worst and were rejected in most cases. The two most favoured models for 1% VaR forecasting were the simple CAViaR models IG-Bayesian and RV-Bayesian. These were favoured by the statistical back-tests (that is, they were not rejected in almost all series), as well as by most of the informal criteria across the series. The two most favoured models for 5% VaR forecasting were the two range-based CAViaR models: TRV-Bayesian and RV-Bayesian. When incorporating range information and employing the Bayesian approach, CAViaR models are competitive at worst, and far more accurate at best, relative to a range of popular and well known VaR methods.

Many additional questions emerge for future research. Are tail risk measures during crisis periods as predictable using similar time series models for longer horizons (for example 10-day forecasts)? Due to space limitations, we only focus on 1-day forecasting here. Extensions to include different state variables in the information set, more than two regimes, and a smooth transition function are potential directions for further research.

Table 9
VaR prediction performance using 17 model specifications and the 400 forecasts of the exchange returns.

$\alpha = 1\%$		Violations	Zone	VRate/α	Penalty	Daily capital charge	Quantile criterion function
Euro vs. US	ST	19	Red	4.75	1.000	5.508	11.074
	LT	4	Green	1.00	0.000	5.108	8.358
	RiskMetrics	5	Green	1.25	0.000	4.756	7.242
	GARCH-n	4	Green	$\boxed{1.00}$	0.000	4.816	7.314
	GARCH-t	2	Green	0.50	0.000	5.054	7.436
	SAV-Bayesian	4	Green	$\boxed{1.00}$	0.000	4.769	7.157
	SAV-E&M	5	Green	1.25	0.000	4.766	7.266
	AS-Bayesian	4	Green	$\boxed{1.00}$	0.000	4.827	$\boxed{7.002}$
	AS-E&M	4	Green	$\boxed{1.00}$	0.000	4.819	7.024
	IG-Bayesian	6	Green	1.50	0.000	4.560	7.474
	TCAV-Bayesian	4	Green	$\boxed{1.00}$	0.000	4.828	7.611
	TCAV-E&M	4	Green	$\boxed{1.00}$	0.000	4.916	7.105
	TIG-Bayesian	6	Green	1.50	0.000	4.662	7.583
	RV-Bayesian	7	Green	1.75	0.000	$\boxed{4.438}$	$\boxed{7.014}$
	RV-E&M	2	Green	0.50	0.000	5.157	7.452
	TRV-Bayesian	6	Green	1.50	0.000	$\boxed{4.517}$	7.082
	TRIG-Bayesian	6	Green	1.50	0.000	$\boxed{4.502}$	$\boxed{7.000}$
JP vs. US	ST	11	Yellow	2.75	0.637	5.067	10.364
	LT	2	Green	0.50	0.000	5.399	$\boxed{8.163}$
	RiskMetrics	6	Green	1.50	0.000	$\boxed{4.696}$	8.882
	GARCH-n	4	Green	$\boxed{1.00}$	0.000	$\boxed{4.780}$	8.222
	GARCH-t	3	Green	0.75	0.000	$\boxed{5.305}$	$\boxed{8.206}$
	SAV-Bayesian	3	Green	0.75	0.000	5.567	8.311
	SAV-E&M	5	Green	1.25	0.000	5.101	9.643
	AS-Bayesian	3	Green	0.75	0.000	5.939	8.730
	AS-E&M	4	Green	$\boxed{1.00}$	0.000	5.852	8.933
	IG-Bayesian	3	Green	0.75	0.000	5.561	8.369
	TCAV-Bayesian	3	Green	0.75	0.000	5.860	9.047
	TCAV-E&M	4	Green	1.00	0.000	5.649	10.568
	TIG-Bayesian	3	Green	0.75	0.000	5.812	$\boxed{8.098}$
	RV-Bayesian	3	Green	0.75	0.000	6.021	8.448
	RV-E&M	5	Green	1.25	0.000	5.240	9.985
	TRV-Bayesian	3	Green	0.75	0.000	5.809	8.336
	TRIG-Bayesian	3	Green	0.75	0.000	5.749	8.263

Table 10
VaR prediction performance using 17 model specifications and the 400 forecasts of the exchange returns.

$\alpha = 5\%$	Euro vs. US			JP vs. US		
	Violations	VRate/α	Quantile criterion function	Violations	VRate/α	Quantile criterion function
ST	28	1.40	30.873	24	1.20	30.350
LT	22	1.10	29.164	20	$\boxed{1.00}$	$\boxed{27.933}$
RiskMetrics	22	1.10	28.226	23	1.15	28.957
GARCH-n	22	1.10	28.271	19	0.95	28.362
GARCH-t	24	1.20	28.455	19	0.95	28.187
SAV-Bayesian	19	0.95	28.154	19	0.95	28.413
SAV-E&M	19	0.95	28.118	19	0.95	28.378
AS-Bayesian	20	1.00	28.254	18	0.90	28.855
AS-E&M	20	1.00	28.241	17	0.85	$\boxed{28.818}$
IG-Bayesian	24	1.20	28.302	19	0.95	28.068
TCAV-Bayesian	18	0.90	28.483	19	0.95	29.462
TCAV-E&M	16	0.80	28.363	23	1.15	28.908
TIG-Bayesian	22	1.10	$\boxed{27.910}$	19	0.95	28.403
RV-Bayesian	22	1.10	$\boxed{28.017}$	19	0.95	$\boxed{28.412}$
RV-E&M	15	0.75	28.596	18	0.90	28.447
TRV-Bayesian	19	$\boxed{0.95}$	28.396	18	0.90	29.142
TRIG-Bayesian	21	$\boxed{1.05}$	28.454	20	$\boxed{1.00}$	29.324

Table 11
Counts of model rejections for the three quantile tests across each market, the portfolio and two exchange rate series.

Specification	1%				5%			
	LR_{uc}	LR_{cc}	DQ	Total	LR_{uc}	LR_{cc}	DQ	Total
ST	8	8	8	8	6	7	8	8
LT	5	5	7	7	0	0	6	6
RiskMetrics	4	4	5	5	1	0	5	5
GARCH-n	4	4	5	5	1	1	5	5
GARCH-t	1	0	5	5	2	1	6	6
SAV-Bayesian	1	0	3	3	0	0	4	4
SAV-E&M	4	3	5	5	2	1	5	5
AS-Bayesian	1	1	3	3	0	0	5	5
AS-E&M	5	5	5	5	3	1	6	6
IG-Bayesian	0	0	1	$\boxed{1}$	0	0	2	$\boxed{2}$
IG-E&M	2	1	4	4	0	1	4	4
TCAV-Bayesian	4	2	4	4	0	0	4	4
TIG-Bayesian	2	1	4	4	0	0	5	5
RV-Bayesian	0	0	1	$\boxed{1}$	2	0	2	$\boxed{2}$
RV-E&M	2	1	4	4	2	2	4	4
TRV-Bayesian	1	0	4	4	1	0	3	3
TRIG-Bayesian	1	1	3	3	2	0	4	4

Table 12
Summary statistics for the VaR forecast criteria in Tables 4 and 5.

1%	VRate/α			MRC		Quantile criterion	
	Mean	Median	RMSD	Mean	Median	Mean	Median
ST	**4.656**	**4.710**	**3.286**	12.788	14.907	**29.599**	**34.192**
LT	1.974	2.330	1.160	**14.295**	**16.442**	28.350	31.432
RiskMetrics	2.111	2.110	1.256	13.062	14.972	23.347	27.040
GARCH-n	1.954	2.000	1.170	12.688	14.421	22.712	26.778
GARCH-t	$\boxed{1.406}$	1.445	0.754	12.904	14.423	22.093	25.512
SAV-Bayesian	1.496	1.555	0.677	12.792	14.620	21.858	$\boxed{24.872}$
SAV-E&M	2.053	1.890	1.308	13.048	14.654	23.128	25.602
AS-Bayesian	1.549	1.765	0.810	12.554	14.488	22.128	$\boxed{24.531}$
AS-E&M	2.239	2.890	1.612	12.931	14.669	24.088	26.410
IG-Bayesian	1.473	1.530	$\boxed{0.574}$	$\boxed{12.436}$	$\boxed{14.395}$	$\boxed{21.283}$	24.996
IG-E&M	1.603	1.540	0.793	12.861	14.799	$\boxed{21.631}$	24.987
TCAV-Bayesian	1.829	1.985	0.973	12.924	15.412	22.871	$\boxed{24.369}$
TIG-Bayesian	1.753	1.890	0.848	12.734	14.802	22.128	25.644
RV-Bayesian	$\boxed{1.269}$	$\boxed{1.225}$	$\boxed{0.601}$	$\boxed{12.469}$	$\boxed{14.277}$	$\boxed{21.321}$	25.078
RV-E&M	1.656	$\boxed{1.375}$	1.148	$\boxed{12.131}$	$\boxed{13.852}$	22.118	25.298
TRV-Bayesian	1.430	1.415	$\boxed{0.664}$	12.702	14.570	22.028	26.653
TRIG-Bayesian	$\boxed{1.378}$	$\boxed{1.330}$	0.689	12.741	14.676	22.435	26.815
5%							
ST	**1.696**	**1.755**	**0.791**			83.535	101.520
LT	$\boxed{1.066}$	$\boxed{1.075}$	$\boxed{0.138}$			**89.503**	**104.950**
RiskMetrics	1.244	1.29	0.307			80.308	96.915
GARCH-n	1.248	1.24	0.327			80.341	97.492
GARCH-t	1.306	1.31	0.384			80.487	97.563
SAV-Bayesian	1.168	1.2	$\boxed{0.241}$			80.803	97.653
SAV-E&M	1.224	1.285	0.336			80.669	97.541
AS-Bayesian	1.215	1.245	0.291			79.225	97.340
AS-E&M	1.274	1.32	0.429			80.108	97.864
IG-Bayesian	1.185	1.2	$\boxed{0.241}$			80.335	97.826
IG-E&M	1.166	1.18	0.258			80.024	97.234
TCAV-Bayesian	1.180	1.22	0.256			79.121	97.403
TIG-Bayesian	1.181	1.2	0.249			79.260	97.353
RV-Bayesian	$\boxed{1.139}$	$\boxed{1.06}$	0.258			$\boxed{77.814}$	$\boxed{94.308}$
RV-E&M	1.173	$\boxed{1.085}$	0.406			$\boxed{78.904}$	$\boxed{94.397}$
TRV-Bayesian	$\boxed{1.124}$	$\boxed{1.085}$	$\boxed{0.242}$			78.927	95.640
TRIG-Bayesian	1.194	1.115	0.286			$\boxed{78.876}$	$\boxed{95.402}$

Acknowledgments

The authors thank the Associate Editor and three referees for their insightful comments that helped improve the paper. Cathy Chen is supported by the grants 96-2118-M-002-MY3 and NSC 99-2118-M-035-001-MY2 from the National Science Council (NSC) of Taiwan.

References

Alexander, C. (2008). *Market risk analysis: value-at-risk models*. Chichester, UK: John Wiley & Sons.

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

Basel Committee (1996).

Beckers, S. (1983). Variance of security price return based on high, low and closing prices. *Journal of Business*, *56*, 97–112.

Berkowitz, J., Christofferson, P. F., & Pelletier, D. (2011). Evaluating value-at-risk models with desk-level data. *Management Science*, *57*, 2213–2227.

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

Brandt, M., & Jones, C. (2006). Volatility forecasting with range-based EGARCH models. *Journal of Business and Economic Statistics*, *24*, 470–486.

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., & Lin, E. M. H. (2008). Volatility forecasting using threshold heteroskedastic models of the intra-day range. *Computational Statistics and Data Analysis*, *52*, 2990–3010.

Chen, C. W. S., Gerlach, R., Lin, E. M. H., & Lee, W. C. W. (in press) Bayesian forecasting for financial risk management, pre and post the global financial crisis. *Journal of Forecasting*. http://dx.doi.org/10.1002/for.1237.

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

Chou, R. (2005). Forecasting financial volatilities with extreme values: the conditional autoregressive range (CARR) model. *Journal of Money, Credit and Banking*, *37*, 561–582.

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

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*, 987–1007.

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

Gaglianone, W. P., Lima, L. R., Linton, O., & Smith, D. R. (2011). Evaluating value-at-risk models via quantile regression. *Journal of Business and Economic Statistics*, *29*, 150–160.

Gallant, A. R., Hsu, C. T., & Tauchen, G. E. (1999). Calculating volatility diffusions and extracting integrated volatility. *Review of Economics and Statistics*, *81*, 617–631.

Garman, M. B., & Klass, M. J. (1980). On the estimation of price volatility from historical data. *Journal of Business*, *53*, 67–78.

Geraci, M., & Bottai, M. (2007). Quantile regression for longitudinal data using the asymmetric Laplace distribution. *Biostatistics*, *8*, 140–154.

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. Y. C. (2011). Bayesian time-varying quantile forecasting for value-at-risk in financial markets. *Journal of Business and Economic Statistics*, *29*, 481–492.

Giacomini, R., & Komunjer, I. (2005). Evaluation and combination of conditional quantile forecasts. *Journal of Business and Economic Statistics*, *23*, 416–431.

Guidolin, M., & Timmermann, A. (2006). Term structure of risk under alternative econometric specifications. *Journal or Econometrics*, *131*, 285–308.

Haas, M., Mittnik, S., & Paolella, M. (2006). Value-at-Risk prediction: a comparison of alternative strategies. *Journal of Financial Econometrics*, *4*(1), 53–89.

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

Hastings, W. K. (1970). Monte-Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*, 97–109.

Jorion, P. (1996). Risk: measuring the risk in value at risk. *Financial Analysis Journal*, *52*, 47–56.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica*, *1*, 33–50.

Koenker, R., & Machado, J. A. F. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, *94*, 1296–1310.

Kuester, K., Mittnik, S., & Paolella, M. S. (2006). Value-at-risk prediction: a comparison of alternative strategies. *Journal of Financial Econometrics*, *4*, 53–89.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, *2*, 173–184.

Mandelbrot, B. (1971). When can price be arbitraged efficiently? A limit to the validity of the random walk and martingale models. *Review of Economics and Statistics*, *53*, 225–236.

McAleer, M., & da Veiga, B. (2008b). Single and portfolio models for forecasting value-at-risk thresholds. *Journal of Forecasting*, *27*, 217–235.

McAleer, M., & da Veiga, B. (2008a). Forecasting value-at-risk with a parsimonious portfolio spillover GARCH (PS-GARCH) model. *Journal of Forecasting*, *27*, 1–19.

McAleer, M., Jimenez-Martin, J.-A., & Perez-Amaral, T. (2010a). *Has the Basel II Accord encouraged risk management during the 2008–09 financial crisis?* Available at SSRN: http://ssrn.com/abstract=1397239.

McAleer, M., Jimenez-Martin, J.-A., & Perez-Amaral, T. (2010b). What happened to risk management during the 2008–09 financial crisis? In R. W. Kolb (Ed.), *Lessons from the financial crisis: causes, consequences, and our economic future* (pp. 307–316). New York: Wiley.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., & Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, *21*, 1087–1091.

Morgan, J. P. (1996). *Riskmetrics* (4th edn). J. P. Morgan Technical Document. New York: J. P. Morgan.

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

Tsionas, E. G. (2003). Bayesian quantile inference. *Journal of Statistical Computation and Simulation*, *9*, 659–674.

Yu, P. L. H., Li, W. K., & Jin, S. (2010). On some models for value-at-risk. *Econometric Reviews*, *29*, 622–641.

Yu, K., & Moyeed, R. A. (2001). Bayesian quantile regression. *Statistics and Probability Letters*, *54*, 437–447.

Yu, K., & Zhang, J. (2005). Distribution and applications—a three-parameter asymmetric Laplace distribution and its extension. *Communications in Statistics—Theory and Method*, *34*, 1867–1879.