Evaluation of Conditional Value-at-Risk Models

Tae-Hwy Lee
Department of Economics
University of California
Riverside, CA 92521, U.S.A.
Tel: +1 909 7875037
Fax: +1 909 7875685
Email: tae.lee@ucr.edu

Burak Saltoğlu
Department of Economics
Marmara University
Istanbul, 81040, Turkey
Tel: +90 216 3368487
Fax: +90 216 3464356
Email: saltoglu@marun.edu.tr

February 2002
Very Preliminary & Incomplete

Abstract

We evaluate predictive performance of a selection of conditional and unconditional Value-at-Risk (VaR) models. By filtering various time varying volatility models we predict risk forecasts and compare them with that of unconditional VaR models. We consider VaR models based on ARCH models, historical simulation, Monte Carlo, nonparametric quantile regressions, and extreme value theory (EVT). In addition, we also apply these methods to the standardized residuals to compute what is often called as conditional VaR (CVaR). The return series are standardized using various conditional variance models such as GARCH, FIGARCH, EGARCH models. We apply these methods to the foreign exchange and stock markets of several countries that suffered recent financial crisis, and compare their performances in terms of various evaluation criteria using the tests of White (2000) and Hansen (2001).

Key Words: VaR, CVaR, EWMA, GARCH, FIGARCH, historical simulation, Monte Carlo method, nonparametric quantile regression, extreme value theory, generalized extreme value distribution, generalized Pareto distribution, Hill estimator, predictive ability, bootstrap, CaViaR, emerging market crisis, Hansen's skewed t.

JEL Classification: C3, C5, G0.

1 Introduction

Increasing financial fragility in emerging markets and the extensive use of derivative products in the developed countries can be two distinguishing features of financial world over the last decade. Consequently, effective use of risk measurement tools has been suggested as a main panacea for mitigating these problems. Uniform methodology of risk measurement called Value-at-Risk (VaR) has received a great attention both from regulatory and academic fronts. During a short span of time, a serious number of papers have studied various aspects of VaR methodology. Concerns are mainly on which conditional distributions or conditional variance models are to be used in the computation. We compare various models in several dimensions.

This paper aims to evaluate predictive performance of various conditional and unconditional Value-at-Risk (VaR) models. This study can also be considered as a stress testing exercise since it covers the recent crises in Southeast Asia, Russia, Brazil and most recently Turkey. We analyzed the daily predictive performance of many VaR models such as exponentially weighted moving average (EWMA), historical simulation (HS), Monte Carlo (MC), bootstrap methods, nonparametric quantile regression (NPQ), extreme value theory (EVT), and conditional autoregressive VaR (CaViaR). These models will be called unconditional VaR models or unfiltered VaR models. We also apply these methods to the standardized residuals by various conditional variance models to compute the VaR, which we call here the conditional VaR (CVaR) models or filtered VaR models. For example, the CVaR models we consider are the Filtered HS, Filtered MC, Filtered bootstrap, Filtered NPQ, Filtered EVT, and Filtered CaViaR models. We compare and evaluate them in several dimensions: (1) VaR versus CVaR models, (2) CVaR models based on various filtering schemes (e.g., GARCH versus FIGARCH), (3) Gaussian distribution versus various skewed and leptokurtic distributions, (4) conventional distributions versus EVT distributions, and (5) parametric versus nonparametric methods. Our forecast evaluation uses one step ahead prediction for 250 day forecast samples to be compatible with the backtesting regulation of Bank for International Settlements (BIS). Various objective functions and evaluation methods are used. Results indicate that the choice of filtering scheme and the degree of severity of crisis appear to be crucial factors in assessing the performance of risk forecasts.

First, conventional VaR models assumes Gaussian normal or its extension to take care of the conditional skewness and kurtosis. These models however have been criticized after the financial turmoil hit Southeast Asia where conventional models are claimed to be less than adequate. Alternatively, models based on extreme value distributions are claimed to perform better during crisis periods. Even though there are some empirical supports for the use of extreme value theory (EVT) (e.g., Pownall and Koedijk 1999, Danielsson and de Vries 1997, Neftçi 2000, Longin 1996, among others), certain problems in EVT methodology have also been documented (e.g., Diebold *et al.,* 2000). In this paper, we intend to compare these two approaches.

Second, one major limitation of the EVT models is the IID assumption. A simple way to go around this assumption is to use a GARCH type filtering scheme to apply the EVT to the standardized residuals, as done by McNeil and Frey (2000). Some researchers have doubts on the use this filtering method over regular EVT models (see Danielsson 2001). Hence, it will be interesting to compare regular EVT versus filtered EVT modeling.

Third, even though using parametric approaches can be very appealing for practical reasons, nonparametric methods may constitute an alternative approach to parametric VaR modeling. This choice is important both in computationally intensive parametric methods and other parametric quantile methods such as variance co-variance methods. For instance, Barone-Adesi *et al.* (2000) show that nonparametric bootstrap instead of Monte Carlo simulation will be more effective in risk modeling. On the other hand, Engle and Manganelli (1999) have developed a recent semi-parametric quantile regression method in the context of VaR, called as conditional autoregressive VaR or CaViaR. A relatively less popular area is to use fully nonparametric quantile method in this framework. For instance, the nonparametric quantile regression (e.g., Cai 2002) may be one alternative. Therefore, we compare parametric and nonparametric methods to estimate the quantiles of a return distribution.

A fourth source of debate in this context is the choice of volatility dynamics. The most conventional volatility model is EWMA of Riskmetrics which is an IGARCH process. Recent evidence in volatility modeling indicates that there is long range dependence in volatility dynamics which makes use of the recent fractionally differenced volatility models developed by Bollerslev and Mikellsen (1996). The choice of volatility is important to estimate the conditional volatilities and for choosing the right model for filtering.

A final categorical discussion can be considered as the recent studies on incorporating the higher moment dynamics such as skewness and kurtosis into risk measurement. Many recent papers (e.g., Hansen 1994, Harvey and Siddique 1999, 2000, and Mittnik and Paolella 2000, among others) suggest that a time varying higher moments will be a better alternative to model volatility and hence better results will be expected for risk measurement. This choice is again useful in modeling the quantiles of a return distribution as well as in the context of filtering return distributions.

However, the recent research in this field has progressed so rapidly that comparing and contrasting the relative performance of the recent methodologies has not been matched. These relative comparison is so vital for risk practitioners and regulators where implementing a better technique may increase risk modeling efficiency tremendously. The major aim of our paper is to compare and contrast various methods and approaches by using the five metric mentioned above. In other words, we have attempted to perform a rigorous risk modeling comparison based on the most recent developments in the field of risk modeling. More concretely, as a first step, we will compare conventional models with EVT based VaR models. Second, the choice of parametric or nonparametric distribution in the context of risk management has been investigated on two grounds. In this category we first investigated the choice of parametric versus nonparametric distribution on numerically intensive method of Monte Carlo simulation. In this category we compared the risk forecasts generated from Monte Carlo with that of a bootstrap method developed by Barone-Adesi *et al.* (2000). In the same category the risk modeling performance of parametric and nonparametric quantile methods has also been conducted. To compare the relative performance of parametric versus nonparametric models the CaViaR models of Engle and Manganelli (1999) is contrasted with that of the most recent nonparametric quantile regression of Cai (2002). As a third comparison category we assessed the risk forecast precision of filtered versus regular unfiltered VaR models. In doing this comparison we had three subcategories: First subcategory is to compare the conventional empirical quantile methods versus filtered empirical quantiles. To this end, we have chosen Hull and White (1998) and a method suggested by us. As a second subcategory, we contrasted regular versus filtered EVT methodologies. In this case we have chosen three EVT methodologies, namely, generalized Pareto distribution (see Neftci 2000), the generalized extreme value distribution (Longin 1996, 2000), and the Hill estimators (Hill 1978, Danilesson and deVries 2001). The choice of filtering scheme constituted our third objective: regular GARCH models versus FIGARCH models. As a fourth category we compared the risk forecasts generated from an IGARCH known as EWMA volatility with that of GARCH and FIGARCH. Finally, we made a comparison of conditionally skewed distributions with that of unconditional volatility models. In this comparison we have used the conditionally skewed distributions of Hansen (1994).

To summarize we have several aspects to compare the various models: (i) unconditional vs. conditional VaR (CVaR), (ii) conditional variance models used in filtering for the CVaR models, (iii) conditional distributions (conventional distributions vs EVT, Gaussian normal vs skewed/leptokurtic distributions), and (iv) parametric vs nonparametric models. The comparisons are conducted in terms of the four different loss functions using the recent tests of White (2000) and Hansen (2001). In each of these categories we have compared the forecasts of each model with the best of all other competing models considered in the paper. In this analysis we compare one-step ahead risk-forecasts. We examine foreign exchange series of Thailand, Korea, Russia, Turkey, Brazil. We consider the conditional VaR models using EVT, quantile regressions, and historical simulation methodologies. We have also used other filtering schemes such as FIGARCH and conditionally skewed distributions. The forecast criteria we use are predictive likelihood of quantile regression, loss functions based on tail loss functions, coverage probability, and the minimum required capital (MRC) of Bank for International Settlement (BIS).

The organization of the paper is as follows. In Section 2 we discuss various VaR methodologies. In Section 3 forecast evaluation criteria and the reality check are discussed. Section 4 presents the empirical results.

2 VaR Models

Consider the return series $\{y_t\}_{t=1}^T$ of a financial asset. The Value at Risk, denoted as $VaR_t(\alpha)$, can be defined as the conditional quantile

$$\Pr(y_t \leq VaR_t(\alpha)|\mathcal{F}_{t-1}) = \alpha. \tag{1}$$

To establish some notation, suppose $\{y_t\}_{t=1}^T$ follows the stochastic process

$$y_t = \mu_t + \varepsilon_t, \tag{2}$$

where $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$ and $E(\varepsilon_t^2|\mathcal{F}_{t-1}) = \sigma_t^2$ given the information set $\mathcal{F}_{t-1}$ (σ-field) at time $t-1$. Let $z_t \equiv \varepsilon_t/\sigma_t$ have the conditional distribution Φ_t with zero conditional mean and unit conditional variance, i.e., $z_t|\mathcal{F}_{t-1} \sim \Phi_t(0,1)$. We now turn to various methods of estimating $VaR_t(\alpha)$. Table 1 summarizes all the models we consider in the paper.

2.1 Variance-Based VaR Models

The first method is the most standard approach, often called variance-covariance methods. In this paper because we consider a single stock index instead of portfolio we do not consider covariances and thus it may be called as variance methods. In this method, $VaR_t(\alpha)$ can be estimated by

$$VaR_t(\alpha) = \mu_t + \Phi_t^{-1}(\alpha)\sigma_t. \tag{3}$$

Hence estimation of the VaR involves the estimation of $\Phi_t(\cdot)$, μ_t, and σ_t. We consider various estimation methods of VaR, which may be labeled with different methods of estimating $\Phi_t(\cdot)$ and σ_t.

We either assume a certain parametric distribution for $\Phi_t(\cdot)$ (e.g. normal distribution, Student-t distribution, generalized error distribution (GED), Hansen's (1994) skewed t, and etc.) or estimate it nonparametrically. The conditional distribution $\Phi_t(\cdot)$ is assumed to be constant over time or simply assumed as Gaussian $N(0,1)$ in which case $\Phi_t^{-1}(0.05) = 1.645$ and $\Phi_t^{-1}(0.01) = 2.326$. To take care of the fat tail distributions of financial returns series, it is also often assumed as Student-$t(\nu)$ with ν degrees of freedom. For $t(6)$, $\Phi_t^{-1}(0.05) = 1.943\sqrt{\frac{\nu-2}{\nu}} = 1.586$ and $\Phi_t^{-1}(0.01) = 3.143\sqrt{\frac{\nu-2}{\nu}} = 2.566$. We use $t(6)$ in this paper.

The conditional variance σ_t^2 is estimated with various volatility methods such as a simple moving average model (Alexander 1998), an exponentially weighted moving average (EWMA) model of Riskmetrics, and ARCH models of Engle (1982), Bollerslev (1986), Nelson (1991), and Glosten, Jaganathan, and Runkle (1993).

2.1.1 Moving Average (MA)

The simplest method to calculate the VaR is to estimate the volatility of the asset return by historical moving average variance. In this method we estimate the volatility

$$\sigma_t^2 = \frac{1}{m-1}\sum_{j=1}^{m}(y_{t-j} - \hat{\mu}_t^m)^2 \tag{4}$$

where $\hat{\mu}_t^m = \frac{1}{m}\sum_{j=1}^{m} y_{t-j}$. See Alexander (1998) for its empirical advantages and disadvantages. This method will be denoted as MA(m). In our empirical part, we use MA(200).

2.1.2 Exponentially Weighted Moving Average (EWMA)

As is accepted in the empirical finance literature, volatility is changing across time which requires a different model building practice for risk than the standard static estimation methods. The most popular volatility model in this framework is the Riskmetrics model of J.P. Morgan (1995), which is an IGARCH specification of the following form

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)(y_{t-1} - \hat{\mu}_t)^2, \tag{5}$$

where $\hat{\mu}_t = \frac{1}{t-1}\sum_{j=1}^{t-1} y_{t-j}$. We assume $\lambda = 0.94$, as calibrated by J.P. Morgan's Riskmetrics, which substantially reduces the volatility computations.

2.1.3 GARCH Models

The standard GARCH model sets

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2, \tag{6}$$

where L denotes the lag operator. We consider two distributions of z_t, $N(0,1)$ and Hansen's (1994) skewed t, which will be denoted as GARCH_N and GARCH_H.

We also consider the FIGARCH model of Baillie and Mikkelsen (1996). By following Bollerslev and Mikkelsen (1996) we rearrange the above equation to get

$$[1 - \alpha(L) - \beta(L)]\varepsilon_t^2 = \omega + [1 - \beta(L)]v_t. \tag{7}$$

where $v_t \equiv \varepsilon_t^2 - \sigma_t^2$. Furthermore, if we let $\phi(L)(1-L) = 1 - \alpha(L) - \beta(L)$, we can introduce fractionally integrated GARCH (or FIGARCH) models as

$$\phi(L)(1-L)^d\varepsilon_t^2 = \omega + [1 - \beta(L)]v_t. \tag{8}$$

Interestingly, the above equation nests three of the alternative volatility models we will use in this paper, namely, IGARCH or more popularly known as Riskmetrics volatility or EWMA ($d = 1$), GARCH ($d = 0$), and FIGARCH $0 < d < 1$.

2.2 Historical-Simulation Based VaR Models

An alternative approach to VaR estimation is to use nonparametric methods. The most conventional methods of this method is called Historical Simulation (HS). One criticism on these methods is to use volatility updating since the conventional Historical Simulation methodology ignores this. For this purpose we have used a method developed by Hull and White (1998). In addition to this method we extended the method of McNeil and Frey (2000) into Historical Simulation context.

2.2.1 Historical Simulation (HS)

The idea behind historical simulation (HS) approach is to assume historical distribution of returns will remain the same over the next periods and hence the empirical distributions of portfolio returns will be used in estimating VaR. this method uses the empirical quantiles of the historical distribution of return series, $\{y_{t-j}\}_{t=1}^{t-1}$, to estimate $VaR_t(\alpha)$ for a given confidence level α. See Jorion (2000, p. 221) for more details on HS.

2.2.2 Filtered HS (HS-GARCH$_N$)

An alternative way to update volatility is to use standardized residuals as follows. Suppose the return series have the following form:

$$y_t = \mu_t + \sigma_t z_t, \tag{9}$$

where the $\{z_t\}$ are innovations which are strictly white noise and σ_t represents the form of time varying volatility. In obtaining a more precise risk estimate by using the empirical quantiles, it is important to take into account of the changing volatility pattern. By following the suggestion of McNeil and Frey (2000) approach it may be more convenient to estimate the empirical quantile on the basis of z_t. Therefore, we can follow a two step procedure as applied in McNeil and Frey (2000). In the first step we estimate the μ_t and σ_t and hence z_t. In the next stage we can estimate a one-step ahead VaR forecasts on the basis of standardized residuals $\{z_t\}$. Formally, the VaR of standardized returns can be estimated by the following formula:

$$\Pr(z_t \leq VaR_t^z(\alpha)|\mathcal{F}_{t-1}) = \alpha. \tag{10}$$

where $VaR_t^z(\alpha)$ is the VaR estimator $\{z_t\}$. Once the empirical quantiles are estimated for standardized residuals, we can estimate one-step predictive distributions of the return series as

$$VaR_t(\alpha) = \mu_t + VaR^z(\alpha)\sigma_t. \tag{11}$$

In other words, once we estimate the VaR of standardized residuals we can transform to get back the VaR of return series. This methodology is similar to the approach adopted by McNeil and Frey (2000) on EVT's. Here we extend this approach to estimate Historical Simulation. In this way we incorporate the time varying volatility pattern into the VaR forecasts.

2.2.3 Filtered Bootstrap HS (Bootstrap HS-GARCH$_N$)

As the choice of parametric form is crucial in modeling risk bootstrap offers an alternative to other methods. This approach was initially discussed in See Jorion (2000, pp. 296-298) or Barone-Adesi *et al.* (2000). This procedure is conducted by sampling from return data $\{y_t\}$. If we want to generate N returns into the future we sample (with replacement) from M returns. Then by randomly picking one return at a time we can project the return series. For instance, if we define $y_{(i)}$ as the ith random return we can constitute the first next day return by

$$S_{t+i} = S_t(1 + y_{(i)}) \tag{12}$$

$i = 1, 2, .., N$. In other words, if we repeat the simulation by N time we may generate a sample path for return series as $S_{t+1}, S_{t+n}, \ldots, S_{t+N}$. Then applying the same principle of Monte Carlo or HS methods we can obtain VaR forecasts. The advantage of the above approach over Monte Carlo is that this method does not rely on parametric distributional assumptions. The only deficiency of this method is the IID assumption of return distributions. In this case, as suggested by Jorion (2000) and studied by Barone-Adesi *et al.* (2000) a filtering method is suggested. In this method we bootstrap from standardized residuals $\{z_t\}$ and then for each sample realization we apply a simple transformation to incorporate the effect of time varying volatility (similar to the one described in HS method). After incorporating the volatility, we use the following recursion to get the sample path of the levels via filtered bootstrap. If we denote the first filtered return series as $y^*_{(1)}$ we can write down the recursion of spot prices

$$S_{t+i} = S_t(1 + y^*_{(i)}) \tag{13}$$

where

$$y^*_{(i)} = \mu_{t+1} + z_{(i)}\sigma_{t+1}. \tag{14}$$

$i = 1, 2, \ldots, N$ and μ_{t+1}, σ_{t+1} are the one step ahead forecasts of the mean returns and volatilities respectively. Hence, the dependency of the random return series are taken into account via the above equation. We filter the volatility by using a time varying volatility method which is mostly a GARCH process. Then, by repeating the above recursion for N times we will obtain a sample path of spot prices. To obtain the VaR from this path we use the similar procedure applied in Monte Carlo method. In this study we have used GARCH as our volatility filtering scheme and choose $M = 250$ and $N = 250$. We choose 1000 bootstrap trials.

2.2.4 Hull and White's Modified HS (HWHS)

As explained before one of the major deficiencies of regular HS is to assume the return distributions do not change over time. A paper by Hull and White (1998) suggested that HS can be improved by taking into account of the volatility changes experienced during the period covered by the historical data. In order to reflect the current market volatility condition they used the ratio of current volatility to the daily volatility at the time of the observation. In other words, if the current volatility of a return series is higher than a month ago the data observed a month ago understates the magnitude of risk. On the other hand, if the current volatility is lesser than that of its past history the reverse is true. To overcome these problems, the ratio of current and historical volatilities are suggested to be incorporate to the historical observations. In other words, in this method the historical observations from which the empirical quantiles will be selected, are adjusted such that the new series reflect the relative volatility changes. Formally, volatility updated return series are defined to be

$$y_t^* = y_t \frac{\sigma_N}{\sigma_t} \tag{15}$$

where σ_N is the most recent volatility estimate σ_t is the historical estimate of the daily volatility, y_t is the regular return series, and y_t^* is the transformed return series. Then the HS will be estimated on the basis of y_t^* rather than y_t. In this way we can reflect the changes in the volatility pattern across time in risk modeling. In our empirical application we have chosen GARCH(1,1) as our volatility model and adjusted return series by applying the formula in (15).

2.3 Simulation Based VaR Models

There are various numerical methods employed in VaR methodology. An underlying stochastic process assumed to govern the dynamics of the asset prices is used to simulate the future values of the asset. One of the most popular stochastic process in the asset pricing context is the geometric Brownian motion (GBM) given as

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t \tag{16}$$

where W_t is a standard Wiener process, and μ_t and σ_t are the drift and the volatility parameters, respectively. This is a rare example of an explicitly solvable stochastic differential equation with its solution being

$$S_t = S_0 \exp\left([\mu_t - \frac{1}{2}\sigma_t^2]t + \sigma_t W_t \right).$$

See Broadie and Glasserman (1998) for more discussion. Thus, simulating values of S_t reduces to simulating values of W_t. where ΔW_t is zero mean, unit variance random error term, and μ_t and σ_t are drift and volatility parameters in discrete time. We simulate S_t in two ways:

2.3.1 Monte Carlo Simulation

Based on random numbers ΔW_t generated from the standard normal distribution. One can then simulate the sample paths of $\{S_t\}$. As it has been used in practice with $\mu_t = \mu$ and $\sigma_t = \sigma$, we implement it in two ways, with constant μ and σ and with time varying μ_t and σ_t. The former will be denoted as MC GBM, and the latter as MC GBM-GARCH$_N$ when σ_t^2 is estimated from GARCH$_N$. See Table 1.

2.3.2 Bootstrap Simulation

Alternatively, one can bootstrap ΔW_t with estimated μ_t and σ_t. The method will be denoted as Bootstrap GBM and Bootstrap GBM-GARCH$_N$.

2.4 Quantile Regressions Based VaR Models

VaR, which is basically a quantile estimator, can be estimated within the context of non-parametric quantile regression methods.

2.4.1 Nonparametric Quantile Regression (NPQ)

There has been various studies made on parametric quantile estimation where a classic reference is Koenker and Bassett (1978). On nonparametric quantile estimation, recent references are Samanta (1989), Lejeune and Sarda (1988), and Xiang (1996) where the most recent nonparemetric quantile method is developed by Cai (2002). Cai's method is in fact based on the weighted Nadaraya-Watson kernel regression of Hall *et al.* (1999).

Let $Z = (Y, X')'$ be a stationary random vector with the joint density $f(y, x)$ and the cumulative distribution function (CDF) $F(y, x)$. Then the conditional CDF is

$$F(y|x) = \frac{\int_{-\infty}^{y} f(\eta, x) d\eta}{F_X(x)}, \tag{17}$$

where $F_X(x)$ is the marginal CDF of X. Hence the quantile function can be found as

$$q_\alpha = \inf\{y \in \mathbb{R} | F(y|x) \geq \alpha\}, \tag{18}$$

where $F(y|x)$ is the conditional distribution of Y_t and can be estimated in various ways. The approach suggested by Cai (2002) is to use the weighted Nadaraya-Watson approach.

$$F(y|x) = \frac{\sum_{t=1}^{n} p_t(x) K(\frac{x_t - x}{h}) \mathbf{1}(Y_t \leq y)}{\sum_{t=1}^{n} p_t(x) K(\frac{x_t - x}{h})} \tag{19}$$

where $K(\cdot)$ is the kernel function, h is the bandwidth, $\mathbf{1}(.)$ is the indicator function and p_t is the weight function. The choice of the weight function is based on maximizing the empirical likelihood function defined in Cai (2002). In other words, maximizing $\sum_{t=1}^{n} \log(p_t(x))$ subject to $\sum_{t=1}^{n} p_t(x) = 1$ through Lagrange multiplier $p_t(x)$ takes the following form:

$$p_t(x) = n^{-1} \{1 + \lambda(x_t - x) K_h(x_t - x)\}^{-1} \tag{20}$$

where λ is chosen to maximize

$$L_n(\lambda) = \frac{1}{nh} \sum\nolimits_{t=1}^{n} \log\{1 + \lambda(x_t - x) K_h(x_t - x)\} \tag{21}$$

In our study the optimal bandwidth selection can be made by modifying the standard cross validation method. This can be represented as

$$h_{CV} = \arg\min_h \sum_{t=1}^{n} \rho_\alpha(y_t - q_\alpha^{-t}(h)) \tag{22}$$

where $\rho_\alpha(\cdot)$ is the loss function defined by Koenker and Bassett (1978), and $q_\alpha^{-t}(h) = \inf\{y \in R | F^{-t}(y|x) \geq \alpha\}$ and $F^{-t}(y|x)$ is a leave-one-out estimator of the conditional CDF.

Estimating and forecasting VaR via nonparametric quantile regression is straightforward. For instance to estimate $VaR_t(\alpha) \equiv q_\alpha$, one needs to estimate the conditional CDF by using the kernel method and then estimate corresponding quantiles. We used the Gaussian kernel where Gaussian kernel $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}u^2)$. Optimal bandwidths are chosen by the robustified cross validation process explained before. In Section 4, we use $x_t = y_{t-1}$. The rolling scheme is used for estimation with the estimation sample size of 1250. That is, for each out of sample forecasting point, we have dropped one distant observation and add a recent one. The optimal bandwidth is chosen for each out of sample data point.

2.4.2 Conditional Autoregressive VaR Models (CaViaR)

Engle and Manganelli (2000) suggest that VaR can be estimated by modeling only the quantiles rather than the whole distribution by using the idea similar to the GARCH modeling:

$$VaR_t = a_0 + a_1 VaR_{t-1} + f(x_t|\theta), \tag{23}$$

where $x_t \in \mathcal{F}_{t-1}$, θ is a parameter, $f(\cdot)$ is a function to explain the VaR model. In our study two of the CaViaR specifications are chosen:

Symmetric Absolute Value (CaViaR$_S$)

$$VaR_t = a_0 + a_1 VaR_{t-1} + a_2|y_{t-1}| \tag{24}$$

Asymmetric Slope Value (CaViaR$_A$)

$$VaR_t = a_0 + a_1 VaR_{t-1} + a_2|y_{t-1}| + a_3|y_{t-1}| \cdot \mathbf{1}(y_{t-1} < 0) \tag{25}$$

Then the estimation can be made via quantile regression. Due to the nondifferentiability the estimation can be achieved by linear programming method. More details can be found in Koenker and Basset (1978) and for a recent discussion on quantile methods Buchkinsky (1998). Filtered versions of the NPQ and CaViaR models are also considered. See Table 1 for the summary of the models we consider.

2.5 Extreme Value Theory

The VaR methodologies reviewed so far ignore extreme price movements and directly model the whole return distribution. The main focus of EVT however, is to model the tails rather than the entire return distributions. There has been various theoretical and empirical studies in this field. For instance, Embrehcts *et al.* (1997) give an excellent review of EVT. Longin (1996, 2000) use GEV to estimate the tail index via maximum likelihood estimation. Neftçi (2000) on the other hand uses GPD. We consider the three EVT models, namely, GEV, GPD, and Hill's tail index estimation.

2.5.1 Generalized Extreme Value (GEV) Distribution: Longin (1996, 2000)

Consider the return series $\{y_t\}_{t=1}^T$ of a financial asset and the ordered return series $\{y_{(t)}\}_{t=1}^T$ in increasing order $y_{(t)} \leq y_{(t+1)}$ for all t. The sample minimum is $y_{(1)}$ over T sample period.

If the returns are IID with the CDF $F_Y(y)$, then the CDF of the minimal return, denoted by $G_Y(y)$, is given by

$$\begin{aligned} G_Y(y) &= \Pr(y_{(1)} \le y) = 1 - \Pr(y_{(1)} > y) = 1 - \prod_{t=1}^{T} \Pr(y_t > y) \\ &= 1 - \prod_{t=1}^{T} [1 - \Pr(y_t \le y)] = 1 - [1 - F_Y(y)]^T. \end{aligned} \tag{26}$$

Thus $G(y)$ is degenerated as $T \to \infty$. Hence, we seek a limit law $H_X(x)$ with which a normalization $x_T = (y_{(1)} - \beta_T)/\delta_T$ does not degenerate as $T \to \infty$ for suitable normalizing constants β_T and $\delta_T > 0$. The limiting distribution of x_T is the generalized extreme value (GEV) distribution of Mises (1936) and Jenkinson (1955) of the form

$$H_X(x) = 1 - \exp(-(1 + \tau x)^{\frac{1}{\tau}}) \tag{27}$$

for $1 + \tau x > 0$. The corresponding limiting density function of $\{x_T\}$ as $T \to \infty$, obtained by differentiating $H_X(x)$, is given by

$$h_X(x) = (1 + \tau x)^{\frac{1}{\tau} - 1} \exp(-(1 + \tau x)^{\frac{1}{\tau}}). \tag{28}$$

Hence the approximate density of $y_{(1)}$ for given T may be obtained by change of variables which is

$$h_Y(x_T) = \frac{1}{\delta_T}(1 + \tau x_T)^{\frac{1}{\tau} - 1} \exp(-(1 + \tau x_T)^{\frac{1}{\tau}}), \tag{29}$$

where $\frac{1}{\delta_T}$ is the Jacobian of the transformation.

Hence the three parameters $\theta_T = (\tau\ \beta_T\ \delta_T)'$ may be estimated by maximum likelihood method. To implement it, Longin (1996, 2000) partition T samples into m non-overlapping subsamples each with n observations. In other words, if $T = mn$, the ith subsample of the return series is $\{y_{(i-1)n+j}\}_{j=1}^{n}$ for $i = 1, \ldots, m$. If $T < mn$, we drop some observation in the first subsample so that it has less than n observations. The collection of the subperiod minima is then $\{y_{n,i}\}$ where $y_{n,i} = \min_{1 \le j \le n}\{y_{(i-1)n+j}\}$, $i = 1, \ldots, m$. The likelihood function of the subperiod minima is

$$\prod_{i=1}^{m} h_Y(x_n) = \prod_{i=1}^{m} h_Y\left(\frac{y_{n,i} - \beta_n^i}{\delta_n^i}\right). \tag{30}$$

Assuming $\theta_n^i = \theta_n$ for all subperiods $i = 1, \ldots, m$, θ_n can be estimated for a numerical optimization of the (log) likelihood.

Consider the probability that the subperiod minimum $y_{n,i}$ is less than y_n^* under the limit law (27). Denoting $x_n^* = \frac{y_n^* - \beta_n}{\delta_n}$, it is

$$H_X(x_n^*) = H_X\left(\frac{y_n^* - \beta_n}{\delta_n}\right) = \Pr\left(\frac{y_{n,i} - \beta_n}{\delta_n} \leq \frac{y_n^* - \beta_n}{\delta_n}\right) = \Pr(y_{n,i} \leq y_n^*), \tag{31}$$

which is therefore equal to

$$G_Y(y_n^*) = 1 - [1 - F_Y(y_n^*)]^n = 1 - (1-\alpha)^n, \tag{32}$$

where the second equality holds if $y_n^* = VaR(\alpha)$. Hence, equating (31) and (32), we get

$$H_X(x_n^*) = 1 - \exp\left(-(1 + \tau x_n^*)^{\frac{1}{\tau}}\right) = 1 - (1-\alpha)^n \tag{33}$$

which yields the VaR forecasts

$$y_n^* = VaR(\alpha) = \beta_n - \frac{\delta_n}{\tau}\left\{1 - [-\ln(1-\alpha)^n]^\tau\right\}. \tag{34}$$

We denote this method as Longin(n) where n is the size of the subperiod. In Section 4, we use $n = 10$ and 20. Tsay (2000) provides an excellent exposition of this method and other VaR models.

2.5.2 Generalized Pareto Distribution (GPD): Neftci (2000)

An alternative approach to GEV method is based on *exceedances over threshold* (Smith 1989, Davison and Smith 1990). According to this approach, we fix some high threshold u and look at all exceedances e over u. The distribution of excess values is given by

$$F_u(e) = \Pr(X < u + e | X > u) = \frac{F(u+e) - F(u)}{1 - F(u)}, \quad e > 0. \tag{35}$$

Pickands (1975) shows that the asymptotic form of $F_u(e)$ is

$$H(e) = 1 - \left(1 - \frac{\tau e}{\delta}\right)^{1/\tau}, \tag{36}$$

where $\delta > 0$ and $1 - \frac{\tau e}{\delta} > 0$. This is known as the generalized Pareto distribution (GPD) with its density

$$h(e) = \frac{1}{\delta}\left(1 - \frac{\tau e}{\delta}\right)^{1/\tau - 1}. \tag{37}$$

Let $\{e_i\}_{i=1}^n$ be the sample of exceedances over threshold with its sample size n. The likelihood $\prod_{i=1}^n h(e_i)$ may be maximized to estimate $\theta = (\tau\ \delta)'$. Once $\hat{\theta} = (\hat{\tau}\ \hat{\delta})'$ is estimated $VaR(\alpha)$ can be estimated as follows. From (35) and (36), we get

$$[1 - F(u+e)] = [1 - F(u)][1 - H(e)]. \tag{38}$$

From this, by letting $[1 - F(u + e)] = \alpha$, $[1 - F(u)] = n/T$, and using the GPD distribution $H(VaR_t(\alpha))$ in (36), we obtain the VaR estimate

$$VaR_t(\alpha) = -\frac{\hat{\delta}}{\hat{\tau}}\left(1 - \left(\frac{T\alpha}{n}\right)^{\hat{\tau}}\right), \tag{39}$$

where T is the total observations and n is the number of exceedances.

The above discussion is for the upper tail (short position). In this paper we focus on lower tails. However, if one uses negative return series $\{-y_t\}_{t=1}^T$ for the variable X in (35), the above discussion continues to apply to the lower tails (long position). Let $x_t = -y_t$. In choosing the threshold value u we follow Neftçi (2000): $u = 1.176 \times \hat{\sigma}_T$ where $\hat{\sigma}_T$ is the standard deviation of $\{x_t\}_{t=1}^T$ from the whole sample and $1.176 = \Phi_t^{-1}(0.10) = 1.440\sqrt{\frac{\nu-2}{\nu}}$ with $t(6)$ distribution being assumed. Therefore, the extreme observations (exceedances over the thresholds) would belong to 10% tails if its true distribution is indeed $t(6)$. The number of $\{x_t\}_{t=1}^T$ that exceeds u is n.

2.5.3 Hill Estimator

As before, we assume the return series $\{y_t\}_{t=1}^T$ are i.i.d. and denote the ordered return series as $\{y_{(t)}\}_{t=1}^T$ in increasing order. Suppose $y_{(n)} < 0$ and $y_{(n+1)} > 0$ so that n is the number of negative returns in the T observations. The GEV distribution in (27) with $\tau < 0$ is known as the Fréchet distribution with the CDF $F_Y(y) = \exp(-y^{\frac{1}{\tau}})$, $y < 0$. As shown in Embrechts *et al.* (1997, p. 321), it reduces to

$$F_Y(y) = 1 - Cy^{\frac{1}{\tau}}, \ |y| \geq u \geq 0 \tag{40}$$

where $C = u^{-1/\tau}$ is a slowly varying function with u being the known threshold. A popular estimator of τ is due to Hill (1975) who shows that its maximum likelihood estimator is

$$\hat{\tau} = -\frac{1}{k}\sum_{t=1}^{k} \ln|y_{(t)}| + \ln|y_{(k+1)}| \tag{41}$$

where $k \equiv k(n) \to \infty$ and $k(n)/n \to 0$. It is known that $\hat{\tau} \to_p \tau$ as $n \to \infty$ (Mason 1982). We choose the sample fraction k using a bootstrap method of Danielsson *et al.* (2000). Once τ is estimated, the VaR estimate can be found from

$$VaR_t(\alpha) = \left[\frac{n}{k}(1-\alpha)\right]^{\hat{\tau}} y_{(k+1)}. \tag{42}$$

See Embrechts *et al.* (1997, p. 347).

2.5.4 Filtered EVT VaR Models

One of the main criticism of the conventional VaR methodology (particularly EVT and HS) is the implicit assumption of IID. In a recent paper McNeil and Frey (2000), a VaR estimation method is proposed based on heteroscedastic return series. After filtering the return series for possible GARCH effect the EVT based VaR model can be fitted by using GPD distribution. In other words, once the conditional volatility estimates for the return series are estimated then return distributions can be standardized. In this section we have employed the same idea for the other EVT models, namely, Hill and Longin methods. If we let the standardized returns $z_t = (y_t - \mu_t)/\sigma_t$. Once the conditional volatility is fitted then one can follow the EVT method for estimating the conditional quantile for z_t. Finally, the Conditional VaR based on EVT can be estimated from $VaR_t(\alpha) = \mu_t + Var_t^z(\alpha)\sigma_t$, as described in the previous section, where $Var_t^z(\alpha)$ is the EVT based quantiles of standardized residuals. McNeil and Frey (2000) have only applied this methodology for the GPD distribution but we extend this framework for the other EVT models namely, GEV and Hill estimators. In other words, for the case of GEV and Hill estimators we first estimated the standardized residuals and then EVT methodology is employed to estimate the quantiles of the standardized residuals. Then this residuals are transformed into conditional VaR by using estimated mean equation and volatility estimates. The critical point here is of course is the choice of volatility model. In addition the estimate of the mean equation plays an important role in this context. In this paper we have chosen two volatility models, GARCH and FIGARCH.

3 Evaluating VaR models

Our primary objective is to compare the various VaR models with the most popular model (the benchmark). We use the EWMA(0.94) as a benchmark model and call it Model 0. When several models using the same data are compared in predictive ability, it is crucial to take into account the dependence among the models. Failing to do so will result in the data-snooping problem which occurs when a model is searched extensively until a match with the given data is found. Conducting inference without taking into account specification search is commonly referred to as "data-snooping" and can be extremely misleading (see Lo and MacKinlay 1999, Ch. 8). White (2000) develops a noble test to compare multiple models in predictive ability accounting for specification search, built on West (1996) and Diebold and

Mariano (1995).

Our evaluation of out-of-sample forecasts proceeds as follows. There are P predictions in all for each model. The first prediction is based on the model with parameters estimated using data from 1 to R, the second on the model with parameters estimated using data from 2 to $R+1, \ldots$, and the last on the model with parameters estimated using data $P-1$ to $R+P-1=T$. Based on the estimated models using a series of rolling samples, each of size R, one-step ahead forecasts are generated for P post-samples, resulting in P forecasts to evaluate each model.

The comparison of l models via given forecast criteria can be formulated as hypothesis testing of some suitable moment conditions of the loss-differential f. Consider an $l \times 1$ vector of moments, $E(f^*)$, where $f^* = f(Z, \beta^*)$ is an $l \times 1$ vector with elements $f_k^* \equiv f_k(Z, \beta^*)$ for a random vector $Z = (Y, X')'$ and $\beta^* \equiv \text{plim } \hat{\beta}_n$. Hypothesis testing for $E(f^*)$ can be conducted whenever the $l \times 1$ sample moment vector

$$\bar{f} = P^{-1} \sum_{t=R}^{T} f(Z_{t+1}, \hat{\beta}_t) \tag{43}$$

has a continuous limiting distribution. For example, when we compare Model k with the benchmark model ($k=0$) using a loss function $Loss$, the kth element of the $l \times 1$ sample moment vector $\bar{f}$ is the loss differential of Model k and Model 0, that is

$$\bar{f}_k = Loss_0 - Loss_k \ \ (k = 1, ..., l). \tag{44}$$

We now define the four forms of the loss function $Loss$.

3.1 Loss Functions

We consider $l = 23$ models plus a benchmark model with total 24 models in Section 4. Let the indicator (a dummy) for the case when return falls beyond the VaR forecast estimated from Model k be denoted as $d_t^k \equiv 1(y_t < VaR_t^k(\alpha))$ for $t = R, \ldots, T$. Let the probability of the unconditional coverage failure be denoted as $p_k^\alpha = \Pr[y_t < VaR_t^k(\alpha)] = \Pr(d_t^k = 1)$. As the indicator $\{d_t^k\}$ has a binomial distribution, the likelihood is $L(p_k^\alpha) = (1-p_k^\alpha)^{n_0}(p_k^\alpha)^{n_1}$ where $n_0 = \sum_{t=R}^T (1-d_t^k)$ and $n_1 = \sum_{t=R}^T d_t^k$ are the number of 0's and 1's in the indicator sequence $\{d_t^k\}_{t=1}^T$. Note that $n_0^k + n_1^k = P$. The indices α and k in d_t^k, p_k^α, $VaR_t^k(\alpha)$, n_0^k, n_1^k will often be suppressed below.

The first loss function A_k^α is based on the quasi log-likelihood (QLL) of $VaR_t^k(\alpha)$, that is

$$QLL_k^\alpha = -P^{-1}\sum_{t=R}^{T} |y_t - VaR_t^k(\alpha)| \cdot \left[\alpha d_t^k + (1-\alpha)(1-d_t^k)\right], \tag{45}$$

which weights the observed deviation from the VaR with the probability with which it is supposed to occur. We define

$$A_k^a \equiv -QLL_k^\alpha. \tag{46}$$

A_k^α is the loss function used in the quantile estimation (see Koenker and Bassett 1978, p. 38, Ferguson 1967, Bertail *et al* 2000). Larger QLL_k^α and thus smaller A_k^α indicate a better goodness of fit.

The second loss function B_k^a is based on the tail mean return (TMR), which is defined as

$$TMR_k^\alpha = P^{-1}\sum_{t=R}^{T} \frac{y_t d_t^k}{n_1^k}, \tag{47}$$

where $n_1^k = \sum_{t=R}^{T} d_t^k$ is the number of tail returns to be used in computing TMR_k^α. TMR is the mean return when return falls below the VaR. When $n_1 = 0$, TMR is not defined and thus not reported. The blank cells in tables correspond to such cases. Larger TMR_k^α indicates a more profit and a smaller loss from trades in the tail. As only the long position (lower tail) is considered in this paper $y_t d_t$ and TMR_k^α take negative values. We define

$$B_k^\alpha \equiv -TMR_k^\alpha, \tag{48}$$

so that all of our objective functions are to be minimized. Thus, the model with smaller B_k^α is a better one. It may be noted that TMR is similar to the expected shortfall, $E[y_t|y_t < VaR_t(\alpha)]$, which is proposed by Artzner *et al* (1999) as an alternative risk measure and used in McNeil and Frey (2000). But they are different in that TMR is the simple average of shortfalls while expected shortfall is the average weighted using a tail density (e.g., an extreme value distribution discussed in Section 2). Note that we use TMR for the purpose of model selection instead of using it as an alternative risk measure.

The third loss function C_k^α is based on the likelihood ratio statistic of Christoffersen (1998). To test whether the probability of the unconditional coverage failure, $p = \Pr[y_t < VaR_t(\alpha)]$, is equal to α, that is to test $H_0 : p = \alpha$ against $H_1 : p \neq \alpha$, consider the indicator d_t which has a binomial distribution with the likelihood $L(p) = (1-p)^{n_0} p^{n_1}$. Under the null, it is $L(a) = (1-\alpha)^{n_0}\alpha^{n_1}$, and thus the likelihood ratio test statistic is

$$LR = -2\ln(L(\alpha)/L(\hat{p})) \xrightarrow{d} \chi(1), \tag{49}$$

where $\hat{p} = \frac{n_1}{n_0+n_1}$ is the maximum likelihood estimator of p. We now define our third loss function based on $LR:$

$$\begin{aligned} C_k^\alpha &\equiv P^{-1}LR \\ &= P^{-1}[-2\ln L(a) + 2\ln L(\hat{p}_k^\alpha)] \\ &= P^{-1}\sum_{t=R}^{T} -2\left[d_t^k \ln(\alpha) + (1-d_t^k)\ln(1-\alpha)\right] + 2\left[d_t^k \ln(\hat{p}_k^\alpha) + (1-d_t^k)\ln(1-\hat{p}_k^\alpha)\right]. \end{aligned} \tag{50}$$

As the smaller LR indicates that the coverage probability p is closer to α, the model with lower C_k^α generates the VaR forecasts with a better coverage probability and thus C_k^α is to be minimized. Note that $C_k^\alpha = 0$ if $\hat{p}_k^\alpha = \alpha$ and $C_k^\alpha > 0$ if $\hat{p}_k^\alpha \neq \alpha$. It should also be noted that the Christoffersen test LR is to test the null hypothesis $H_0 : p_k^\alpha = \alpha$ for a given model k and it is not to compare models, while the White's reality check using C_k^α is for model selection and to compare models. When $n_1^k = 0$, $\hat{p}_k^\alpha = 0$ and thus $\ln(\hat{p}_k^\alpha)$ and C_k^α are not defined. The blank cells in Tables 3-4 correspond to such cases.

The fourth loss function D_k^α is based on BIC's minimum required capital (MRC). Forecast precision of VaR models is also central for Basel Committee's capital requirements (see Jorion pp. 129). In other words, the forecast performance of internal VaR models are used to determine the required capital for commercial banks. For VaR with 99% probability the Basel Committee uses the following capital requirement rule (for $\alpha = 0.99$) for banks to allocate capital (see Jorion p. 65)

$$D_k^\alpha \equiv MRC_k^\alpha \equiv \max\left[m\frac{1}{60}\sum_{i=1}^{60} VaR_{t-i}^k(\alpha), VaR_{t-i}^k(\alpha)\right] \tag{51}$$

where $m = 3$ if $d_t \leq 4$; $m = 3.40$ if $d_t = 5$; $m = 3.50$ if $d_t = 6$; $m = 3.65$ if $d_t = 7$; $m = 3.75$ if $d_t = 8$; $m = 3.85$ if $d_t = 9$; and $m = 4.00$ if $d_t \geq 10$. Therefore our final loss function is a real loss function for a bank where the worse the internal model used the higher the capital will be required to put aside. To put it differently, the better the precision of the bank's internal risk model the lesser penalty it will face. Therefore, the lower the MRC the better the risk precision of a model. It is also important to note that there is no loss function corresponding to VaR with probability of $\alpha = 0.95$. For this case, no m values are assigned by the Basel Committee. Furthermore, in our empirical section we employed only a one step ahead forecasts (rather than 5 days or 10 days ahead forecasts) since the backtesting rules of Basle Committee requires only one day holding period i.e. one step ahead forecasts.

All of the four objective functions A_k^α, B_k^α, C_k^α, and $D_k^{0.99}$ for our VaR reality check will be minimized. Note that all four loss functions are expressed as sample moments of the form $P^{-1}\sum_{t=R}^{T} x_t$ with x_t being the respective summands expressed in (45), (47), (50), and (51), so that test statistics can be formulated based on the sample moment vector of loss differentials as in (43).

Inevitably, each loss function has its advantages and disadvantages depending on the need and perspective. For instance, the regulators (i.e. BIS or central banks) would like to put more weights on coverage probabilities therefore the loss function C would be more suitable for them. If, on the other hand, the focus is the difference between the actual return and the risk forecasts, then the loss A may be the most appropriate. If an institution wants to ensure avoiding potential big losses, it will prefer VaR models to be evaluated under the loss B. Therefore, we have used three different loss functions to reflect the notion that different agents may have different objectives to evaluate risk measurement forecasts. Hence, we use all three objectives in our empirical section.

3.2 Reality Check for Predictive Ability

Suppose one-step predictions are to be made for P prediction periods, indexed from R through T, so that $T = R + P - 1$. Here, P and R may increase as the sample size T increases. The first forecast is based on the model parameter estimator $\hat{\beta}_R$, formed using observations 1 though R, the next based on the model parameter estimator $\hat{\beta}_{R+1}$, formed using observations 2 though $R+1$, and so forth, with the final forecast based on the model parameter estimator $\hat{\beta}_T$. Often, model comparison via forecast criteria can be conveniently formulated as hypothesis testing of some suitable moment conditions. Consider an $l \times 1$ vector of moments, $E(f^*)$, where $f^* = f(Z, \beta^*)$ is an $l \times 1$ vector with elements $f_k^* \equiv f_k(Z, \beta^*)$ for a random vector $Z = (Y, X')'$ and $\beta^* \equiv \text{plim}\ \hat{\beta}_T$. As discussed earlier, hypothesis testing for $E(f^*)$ can be conducted whenever the $l \times 1$ sample moment vector $\bar{f} = P^{-1}\sum_{t=R}^{T} f(Z_{t+1}, \hat{\beta}_t)$ has a continuous limiting distribution.

West (1996, Theorem 4.1) shows that under proper regularity conditions,

$$\sqrt{P}(\bar{f} - E(f^*)) \to N(0, -)\ \text{in distribution} \tag{52}$$

as $P \equiv P(T) \to \infty$ when $T \to \infty$, where - is a $l \times l$ matrix

$$- = \lim_{T\to\infty} \text{var}[P^{-\frac{1}{2}} \sum_{t=R}^{T} f(Z_{t+1}, \hat{\beta}_t)], \tag{53}$$

which is a complicated expression as - depends on the estimated parameter $\hat{\beta}_t$. When either

$$F \equiv E(\partial f^* / \partial \beta) = 0 \tag{54}$$

or $P/R \to 0$ as $T \to \infty$, - can be substantially simplified because then - does not depend on the estimated parameter $\hat{\beta}_t$ and

$$- = \lim_{T \to \infty} \text{var}[P^{-\frac{1}{2}} \sum_{t=R}^{T} f(Z_{t+1},\, \beta^*)], \tag{55}$$

which corresponds to West's (1996) Theorem 4.1(a) and the result of Diebold and Mariano (1995). Here, the effect of using $\hat{\beta}_t$ rather than β^* is asymptotically negligible. One can proceed as if β^* were known and were equal to $\hat{\beta}_t$ in period t. However, when $F \neq 0$, as is the case in this paper due to the use of the indicator d_t in our three loss functions, - is unknown and depends on $\hat{\beta}_t$. In this case, although - is not feasible to derive even asymptotically, a bootstrap procedure may be used to obtain the null distribution of the statistic.

When we compare a single model ($l = 1$) with a benchmark we can use Diebold and Mariano's (1995) test and West's (1996) test with an appropriate estimator of - . When we compare multiple forecasting models ($l > 1$) against a given benchmark model, however, sequential use of Diebold and Mariano (1995) and West (1996) tests may result in a data-snooping bias since the test statistics are mutually dependent due to the use of the same data. See Lo and MacKinlay (1999, Ch.8) and White (2000) for more discussions on the biases due to data snooping. To account for possible bias due to data snooping, we use White's (2000) procedure. The appropriate null hypothesis is that the best model is no better than a benchmark, expressed formally as

$$H_0 : \max\nolimits_{1 \leq k \leq l} E(f_k^*) \leq 0. \tag{56}$$

This is a multiple hypothesis, the intersection of the one-sided individual hypotheses $E(f_k^*) \leq 0,\ k = 1, \ldots, l$. The alternative is that H_0 is false, that is, that the best model is superior to the benchmark. White's (2000) test statistic for H_0 in (56) is formed as follows:

$$\bar{V} \equiv \max_{1 \leq k \leq l} \sqrt{P} \bar{f}_k, \tag{57}$$

which converges in distribution to $\max_{1 \leq k \leq l} Z_k$ under H_0, where the limit random vector $Z = (Z_1, \ldots, Z_l)'$ is $N(0, -)$. This null limit distribution is however unknown (due to unknown -) and not feasible to derive even asymptotically. White (2000) suggests to use the stationary bootstrap of Politis and Romano (1994) to obtain the null distribution of $\bar{V}$. This gives appropriate p-values for testing the null hypothesis that the best model has no predictive superiority relative to the benchmark (White, 2000, Corollary 2.4). The p-value is called the "Reality Check p-value" for data snooping. White (2000, Proposition 2.5) also shows that the test's level can be driven to zero at the same time the power approaches to one as $\bar{V}$ diverges at rate $P^{\frac{1}{2}}$ under the alternative. Implementation of the Reality Check bootstrap and an illustrative example can be found in White (2000). See also Sullivan, Timmermann and White (1998, 1999) for applications to the studies of technical trading rules and calendar effects in asset markets.

White (2000, Theorem 2.3 & Corollary 2.4) shows that a sufficient condition for the validity of the stationary bootstrap procedure is $(P/R) \log\log R \to 0$ as $n \to \infty$, no matter whether $E(\partial f^* / \partial \beta) = 0$. White (2000) only considers differentiable ψ (e.g., MSFE). As noted in White (2000, p. 1100), it is possible to extend his procedure to nondifferentiable ψ (e.g., our four loss functions). Checking White's (2000) proof, we can see that when no parameter estimation is involved, White's (2000) procedure is applicable to nondifferentiable ψ. We expect that when parameter estimation is involved, the impact of parameter estimation uncertainty is asymptotically negligible (so that - is as in (4.1)) when P grows at a suitably slower rate than R. In this case, we conjecture that White's (2000) procedure continues to hold for nondifferentiable ψ no matter whether $\partial E(f^*)/\partial \beta = 0$. However, the proof is much involved and has to be pursued in further work. We note that McCracken's (2000) approach will be useful here.

Under appropriate conditions and under the null hypothesis, the distribution of $\sqrt{P}(\bar{f}^* - \bar{f})$ converges to that of $\sqrt{P}(\bar{f} - E(f^*))$ where $\bar{f}^*$ is obtained from the stationary bootstrap. We obtain the empirical quantiles of the statistic

$$\bar{V}^* = \max_{1 \leq k \leq l} \sqrt{P}(\bar{f}_k^* - \bar{f}_k) \tag{58}$$

and, accordingly, the p-value for testing the null hypothesis that the best model has no predictive superiority relative to the benchmark (White, 2000, Corollary 2.4). This p-value is called the "Reality Check p-value" for data snooping.

The inclusion of $\bar{f}_k$ in (58) guarantees that the statistic satisfies the null hypothesis $E(\bar{f}_k^* - \bar{f}_k) = 0$ for all k. This setting makes the null hypothesis the least favorable to the alternative and consequently, it renders a very conservative test. When a highly misspecified model is introduced, the reality check p-value becomes very large and, depending on the variance of $\bar{f}_k$, it may remain large even after the inclusion of better models. Hence, the White's reality check p-value may be considered as an upper bound of the true p-value. Hansen (2001) considered different adjustments to (58) providing a lower bound for the statistic as well as intermediate values that depend of the variance of $\bar{f}_k$. In Hansen (2001) the statistic (58) is modified as

$$\bar{V}^* = \max_{1 \leq k \leq l} \sqrt{P}(\bar{f}_k^* - g(\bar{f}_k)) \tag{59}$$

Different $g(\cdot)$ functions will produce different bootstrap distributions that are compatible with the null hypothesis. If $g(\bar{f}_k) = \max(\bar{f}_k, 0)$, the null hypothesis is the most favorable to the alternative, and the p-value associated with the test statistic under the null will be a lower bound of the true p-value. Hansen (2001) recommends setting $g(\cdot)$ as a function of the variance of $\bar{f}_k$, i.e.

$$g(\bar{f}_k) = \begin{cases} 0 & \text{if } \bar{f}_k \leq -A_k \\ \bar{f}_k & \text{if } \bar{f}_k > -A_k \end{cases} \tag{60}$$

where $A_k = \frac{1}{4} P^{1/4} \sqrt{\mathrm{var}(\bar{f}_k)}$ and the variance can be estimated from the bootstrap resamples.

4 Empirical Results

Daily nominal foreign exchange (FX) rate series are obtained from Datastream for the Asian countries: Thailand baht, Korean won, Russian rouble, Turkish lira, and Brazilian real. Since the crisis FX crisis hit these countries differently, we analyzed various periods that affected these markets. Table 2 presents the summary of the data. For Korea the sample period was between January 12, 1992 to October 1, 1997 including 1262 observations, i.e., $R = 1262$. We have forecasted the FX rate risk of Korean won until October 9, 1998 using $P = 250$. For Thailand baht we analyzed the period between January 1, 1992 and June 30, 1997 as our in sample period $R = 1434$. For out-of-sample risk forecast analysis we use the daily data until June 15, 1998 with $P = 250$ similar time period where begin our analysis. Daily logarithmic returns for FX rates are analyzed. The FX crisis of Brazil has been investigated by using the daily rates between January 1, 1993 and December 31, 1998. The out-of-sample period was until December 17, 1999, again $P = 250$. For the Russian bond market crisis which resulted in a severe devaluation in Russian rouble has been investigated between May 7, 1993 and March 8, 1998, including 1947 observations, i.e., $R = 1947$. For the Turkish FX market crisis took place in 1994 was investigated between the daily Turkish lira rates beginning from January 1, 1990 to April 1, 1994 with $R = 1087$. We analyze the crisis until December 14, 1995 with $P = 250$ and the analysis contains only short position, i.e., we looked at from the right tail of the FX return distribution.

In the Tables 3 and 4 we present the forecast results for Thailand's baht and Brazilian real. Both of these tables report the test statistic values of Christoffersen (1998) coverage probability tests. The results show some similarities between these two countries. For Thailand, if we investigate the VaR forecasts for $\alpha = 0.05$, we noticed that very rare models have adequate performance. For instance, when the actual coverage probability was 0.05, the estimated probability for EWM model was 0.08 and has a Likelihood Ratio test statistic of 4.1 which rejects the probability of actual and estimated probabilities are equal. This is true almost for each risk models except the GARCH filtered Hill method. For this method we observe a LR of unconditional coverages equal to 0.54 where we can not reject the null that the estimated and actual coverages are equal since the critical value of such test is 3.84. The same is true for the test for conditional coverages. In this case conditional coverages of Filtered Hill method again seems to be satisfactory. When we consider the VaR for probability 0.01 we notice a different result. In this case many of the filtered methods seemed to fare better than other regular models. Particularly, Filtered HS, Filtered EVT with GPD, GEV and Hill estimators, and Filtered NPQ are all seemed to have satisfactory probability coverages. Many unfiltered models such as EWMA, MC GBM, GARCH, and unfiltered EVT models fail to capture the actual coverages. Relatively less efficient forecasting precision of CaViaR model, Hull and White (1998) model, filtered bootstrap model were also revealed from this country study. As before, Filtered HS and NPQ models were relatively efficient in capturing the coverage probabilities of risk forecasts. For this country one of the most interesting results is that the Filtered Nonparametric Methods do at least as well as Filtered EVT models. We did the similar analysis for Brazilian Real which also suffered from financial crisis hit Brazil. The results obtained from this country has some similarities and dissimilarities to Thailand case. First, for $\alpha = 0.05$, there are more models that appear to be more satisfactory than that of Thailand's Baht. Conventional EWMA and conventional Variance-covariance model as well as Filtered Monte Carlo and Hill estimators seemed to be successful. The picture changes somewhat differently when we look at risk forecast at the probability of $\alpha = 0.01$. In this circumstances we can see that the most successful one was again filtered methods. Particularly, the filtered Historical Simulation seems to forecast the actual coverages most efficiently. Other filtered methods, especially, Filtered EVT models such as GPD, GEV, and Hill show a significant coverage probabilities. On the other hand, CaViaR, Hull and White and partly Filtered Bootstrapped methods produced relatively inefficient risk forecasts under this test statistics. A general finding up to this point can be stated that filtering is very useful in detecting risk forecasts but this does not have to be done under EVT methods. Most surprisingly, the most conventional risk model of EWMA volatility does a very good job for this exchange rate series. Particularly, filtered HS method produced very efficient coverage probabilities for both countries and probabilities.

5 Conclusions

In this paper we evaluate predictive performance of various conditional and unconditional Value-at-Risk (VaR) models. This study can also be considered as a stress testing exercise since it covers the recent crises in Southeast Asia, Russia, Brazil and most recently Turkey. We analyzed the daily predictive performance of many VaR models such as exponentially weighted moving average (EWMA), historical simulation (HS), Monte Carlo (MC), bootstrap methods, nonparametric quantile regression (NPQ), extreme value theory (EVT), and conditional autoregressive VaR (CaViaR). These models will be called unconditional VaR models or unfiltered VaR models. We also apply these methods to the standardized residuals by various conditional variance models to compute the VaR, which we call here the conditional VaR (CVaR) models or filtered VaR models. For example, the CVaR models we consider are the Filtered HS, Filtered MC, Filtered bootstrap, Filtered NPQ, Filtered EVT, and Filtered CaViaR models. We compare and evaluate them in several dimensions: (1) VaR versus CVaR models, (2) CVaR models based on various filtering schemes (e.g., GARCH versus FIGARCH), (3) Gaussian distribution versus various skewed and leptokurtic distributions, (4) conventional distributions versus EVT distributions, and (5) parametric versus nonparametric methods. Our forecast evaluation uses one step ahead prediction for 250 day forecast samples to be compatible with the backtesting regulation of Bank for International Settlements (BIS). Various objective functions and evaluation methods are used. Results indicate that the choice of filtering scheme and the degree of severity of crisis appear to be crucial factors in assessing the performance of risk forecasts.

References

Abberger, K. (1997), "Quantile Smoothing in Financial Time Series", *Statistical Papers*, 38, 125-148.

Alexander, C. (1998), "Volatility and Correlation: Measurement, Methods and Applications", in *Risk Management and Analysis*, C. Alexander (ed.), Volume 1, Chapter 4, 125-168, John Wiley, New York.

Basle Committee on Banking Supervision (1996), *Ammendment to the Basle Capital Accord to Incorporate Market Risks*, BIS, Basle, Switzerland.

Beim, D.O. and C.W. Calomiris (2001), *Emerging Financial Markets,* McGraw-Hill.

Barone-Adesi, G., K. Giannopoulos, and L. Vosper (1999), "VaR without Correlations for Non-linear Portfolios", *Journal of Futures Markets*, 19, 583-602.

Barone-Adesi, G., K. Giannopoulos, and L. Vosper (2000), "Filtered Historical Simulation Backtest Analysis", City University Business School, London.

Bertail, P., C. Haefke, D.N. Politis, and H. White (2000), "A Subsampling Approach to Estimating the Distribution of Diverging Statistics with Applications to Assessing Financial Market Risks", UCSD Discussion Paper 2000-01.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31, 307-327.

Bollerslev, T., and H.O. Mikkelsen (1996), "Modeling and Pricing Long Memory in Stock Market Volatility", *Journal of Econometrics*, 73, 151-184.

Boudoukh, J., M. Richardson, and R. Whitelaw (1998), "The Best of Both Worlds", *Risk*, 11, 64-67.

Brodie, M. and P. Glasserman (1998), "Simulation for Option Pricing and Risk Management", in *Risk Management and Analysis*, C. Alexander (ed.), Volume 1, Chapter 5, 173-207, John Wiley, New York.

Butler, J.S. and B. Schachter (1998), "Estimating Value at Risk with a Precision Measure by Combining Kernel Estimation with Historical Simulation", *Review of Derivatives Research*, 1, 371-390.

Buchinsky, M. (1998), Recent Advances in Quantile Regression Models: A Practical Guiedlines for Empirical Research, *Journal of Human Resources*, 33, 1, 88-127.

Cai, Z. (2002), "Regression Quantiles for Time Series", *Econometric Theory*, 18, 169-192.

Chernozhukov, V. (2000), "Conditional Value-at-Risk: Aspects of Modeling and Estimation", MIT.

Chernozhukov, V. (2001), "Conditional Extreme and Near-Extremes", MIT.

Christoffersen, P.F (1998), "Evaluating Interval Forecasts", *International Economic Review*, 39(4), 841-864.

Christoffersen, P.F, F.X. Diebold, and T. Shuerman (1998), "Horizon Problems and Extreme Events in Financial Risk Management", *Economic Policy Review*, Federal Reserve Board of New York, October, 109-118.

Christoffersen, P.F, and V. R. Errunza,(2000), "Towards a Global Financial Architecture: Capital Mobility and Risk Management", *Emerging Market Review*, 1, 2-19.

Crnkovic C. and J. Drachman (1995), "A Universal Tool to Discriminate Among Risk Management Techniques", Corporate Risk Management Group, J.P. Morgan.

Crouhy, M., D. Galai, and R. Mark (1998), "The New 1998 Regulatory Framework for Capital Adequacy", in *Risk Management and Analysis*, C. Alexander (ed.), Volume 1, Chapter 1, 1-37, John Wiley, New York.

Corsetti S., P. Pesenti, and N. Roubini (1999), "What Caused the Asian Currency and Financial Crisis", *Japan and World Economy*, 11, 3, 305-373.

Crouhy, M., D. Galai, and R. Mark (1998), "The New 1998 Regulatory Framework for Capital Adequacy", in *Risk Management and Analysis*, C. Alexander (ed.), Volume 1, Chapter 1, 1-37, John Wiley, New York.

Danielsson, J. (2001), "The Emperor Has No Clothes: Limits to Risk Modelling", *Journal of Money Banking and Finance*, forthcoming.

Danielsson, J., L. deHaan, L. Peng, and C.G. deVries (2000), "Using a Bootstrap to Choose the Sample Fraction in Tail Index Estimation", *Journal of Multivariate Analysis*, forthcoming.

Danielsson, J. and C.G. deVries (1997), "Beyond the Sample: Extreme Quantile and Probability Estimation", LSE and Tinbergen Institute.

Danielsson, J. and C.G. deVries (2000), "Value at Risk and Extreme Returns", LSE and Tinbergen Institute.

Danielsson, J. and J. Morimoto (2000), "Forecasting Extreme Financial Risk: A Critical Analysis of Practical Methods for the Japanese Market", IMES Discussion Paper Series-2000-E-8.

Dave, R. and G. Stahl (1997), "On the Accuracy of VaR Estimates Based on the Variance-Covariance Approach", http://www.olsen.ch/research

Davison, A.C. and R.L. Smith (1990), "Models for Exceedances over High Thresholds", *Journal of Royal Statistical Society, B*, 52, 393-442.

Diebold, F.X. and R.S. Mariano (1995), "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, 13, 253-263.

Diebold, F.X. and A. Santomera (1999), "Financial Risk Management in a Volatile Global Environment", *Asia Risk*, December, 35-36.

Diebold, F.X., T. Schuermann, and J. Stroughair (2000), "Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management", *Journal of Risk Finance*, 1, 30-36.

Dornbusch, R. (2001), "A Pimer on Emerging Market Crises", MIT.

Dowd, K. (1998), *Beyond Value at Risk: The New Science of Risk Management*, Wiley, New York.

Embrechts P., C. Klüppelberg, and T. Mikosch (1997), *Modelling Extremal Events for Insurance and Finance*, Springer Verlag, New York.

Engle, R.F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation", *Econometrica*, 50, 987-1008.

Engle, R.F. and S. Magnelli (1999), CaViaR: Conditional Autoregressive Value at Risk by Quantile Regression, NBER,Working paper, 7341.

FRBNY (2000), *Economic Policy Review,* Special Issue: Lessons from Recent Crises in Asian and Other Emerging Markets, Volume 6, Number 3, September, Federal Reserve Bank of New York.

Gencay, R., F. Selcuk, and A. Ulugulyagci (2001), "High Volatility, Thick Tails, and Extreme Value Theory in Value-at-Risk Estimation", University of Winsor.

Glosten, L. R., R. Jaganathan and D. Runkle (1993), "On the Relationship Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks", *Journal of Finance*, 48, 1779-1801.

Hall, P., R.C.L. Wolff and Q. Yao (1999), "Methods for Estimating A Conditional Distribution Function", *Journal of American Statistical Association*, 94, 154-163.

Hansen, B.E. (1994), "Autoregressive Conditional Density Estimation", *International Economic Review*, 35(3), 705-730.

Hansen, P.R. (2001), "An Unbiased and Powerful Test for Superior Predictive Ability", Brown University.

Harvey, C. and A. Siddique (1999), "Autoregressive Conditional Skewness," *Journal of Financial and Quantitative Analysis* 34, 465-488.

Harvey, C. and A. Siddique (2000), "Conditional Skewness in Asset Pricing Tests," *Journal of Finance* 55, 1263-1295.

Hill, B.M. (1975),"A Simple General Approach to Inference about the tail of a Distribution", *Annals of Statistics*, 19, 1547-1569.

Ho, L.C., P. Burridge, J. Cadle, and M. Theobald (2000), "Value at Risk: Applying the Extreme Value Approach to Asian Markets in Recent Financial Turmoil", *Pacific-Basin Finance Journal*, 8, 249-275.

Hull, J. (1997), *Options Futures and Other Derivatives*, New York, Prentice Hall.

Hull, J. and A. White (1998), "Incorporating Volatility Updating into Historical Simulation Method for Value at Risk," *Journal of Risk*, 1, 5-19.

Jenkinson, A.F. (1955), "The Frequency Distribution of the Annual Maximum (or Minimum) Values of Meteorological Elements", *Quarterly Journal of the Royal Meteorological Society*, 81, 145-158.

J.P. Morgan (1995), *Riskmetrics Technical Manual*, 3ed.

Jorion, P. (2000), *Value at Risk*, 2ed., McGraw Hill, New York.

Krugman, P. (1998), "Japan's Bank Bailout: Some Simple Arithmetic", MIT.

Koenker R. and G. Bassett (1978), "Regression Quantiles", *Econometrica*, 46, 1, 33-50.

Lejeune, M. and P. Sarda (1988), "Quantile Regression: A Nonparametric Approach", *Computational Statistics and Data Analysis*, 6, 229-239.

Lo, A. W. and A.C. MacKinlay (1999), *A Non-Random Walk Down Wall Street*, Princeton University Press, Princeton.

Longin, F. (1996), "The Asymptotic Distribution of Extreme Stock Market Returns", *Journal of Business*, 69, 3, 383-408.

Longin, F. (2000), "From Value at Risk to Stress Testing: The Extreme Value Approach", *Journal of Money Banking and Finance*, 24, 1097-1130.

Manganelli, S. and R.F. Engle(2001), "Value at Risk Models in Finance", European Central Bank.

Mason, D.M. (1982), "Law of Large Numbers for Sum of Extreme Values", *Annals of Probabilty*, 10, 754-764.

McCracken, M.W. (2000), "Robust Out-of-sample Inference", *Journal of Econometrics,* 99(2), 195-223.

McNeil, A.J. and R. Frey (2000), "Estimation of Tail-related Risk Measures for Heteroskedastic Financial Time Series: An Extreme Value Approach", *Journal of Empirical Finance,* 7, 271-300.

Mises, R. von (1936), "La Distribution de la plus grande de n valeurs", Reprinted in *Selected Papers II*, American Mathematical Society, Providence, RI (1954), 271-294.

Mittnik, S. and M.S. Paolella (2000), "Conditional Density and Value-at-Risk Prediction of Asian Currency Exchange Rates", *Journal of Forecasting*, 19, 313-333.

Neftçi, S. (2000), "Value at Risk Calculations, Extreme Events, and Tail Estimation", *Journal of Derivatives*, spring, 23-38.

Nelson, D.B. (1991), "Conditional Heteroscedasticity in Asset Returns: A New Approach", *Econometrica*, 59, 2, 347-370.

Politis, D.N. and J.P. Romano (1994), "The Stationary Bootstrap", *Journal of American Statistical Association*, 89, 1303-1313.

Pownall, R. A.J and K. G. Koedijk (1999), "Capturing Downside Risk in Financial Markets: the Case of the Asian Crisis," *Journal of International Money and Finance*, 18, 853-870.

Rockafellar, R.T and S. Uryasev (2001), "Conditional Value-at-Risk for General Loss Distributions", University of Florida, www.ise.ufl.edu/uryasev.

Rockinger, M. and E. Jondeau (2002), "Entropy Densities with an Application to Autoregressive Conditional Skewness and Kurtosis", *Journal of Econometrics*, 106(1), 119-142.

Samanta, M. (1989), "Nonparametric Estimation of Conditional Quantiles", *Statistics and Probability Letters*, 7, 407-412.

Smith, R.L. (1989), "Extreme Value Analysis of Environmental Time Series: An Application to Trend Detection in Ground-Level Ozone", *Statistical Science*, 4, 367-393.

Stulz, R. (2000), "Why Risk Management Is Not a Rocket Science?", *Financial Times, Mastering Risk Series, volume, number and page*.

Sullivan, R., A. Timmermann, and H. White (1999), "Data-snooping, Technical Trading Rule Performance, and the Bootstrap", *Journal of Finance*, 54, 1647-1692.

Sullivan, R., A. Timmermann, and H. White (2001), "Dangers of Data Mining: the Case of Calendar Effects in Stock Returns", *Journal of Econometrics,* 105(1), 249-286.

Taylor, J.W. (1999), "Evaluating Volatility and Interval Forecasts", *Journal of Forecasting*, 18, 111-128.

Tsay, R. (2001), *Analysis of Financial Time Series*, Wiley.

West, K.D. (1996), "Asymptotic Inference about Predictive Ability", *Econometrica*, 64 1067-1084.

White, H. (2000), "A Reality Check for Data Snooping", *Econometrica*, 68, 5, 1097-1126.

Xiang, X. (1996), "A Kernel Estimator of a Conditional Quantiles", *Journal of Multivariate Analysis*, 59, 206-216.

TABLE 1. VaR Models

	Variance Models	Historical Simulation	Simulating GBM	Quantile Regressions	EVT
Unfiltered VaR	MA	HS	MC GBM	NPQ	GEV
	EWMA		Bootstrap GBM	CaViaR_S	GPD
	GARCH_N			CaViaR_A	Hill
	FIGARCH_N				
	GARCH_H				
Filtered VaR		HS-GARCH_N	MC GBM-GARCH_N	NPQ-GARCH_N	GEV-GARCH_N
(Filtered		Bootstrap HS-GARCH_N	Bootstrap GBM-GARCH_N	CaViaR_S-GARCH_N	GPD-GARCH_N
with GARCH_N)		HWHS		CaViaR_A-GARCH_N	Hill-GARCH_N

Notes: (1) The subscripts N and H in GARCH_N, GARCH_H, and FIGARCH_N denote the distribution used for the maximum likelihood estimation. We use Gaussian normal and Hansen's (1994) skewed t, which are denoted as N and H, respectively. For example, GARCH_N denotes the GARCH model estimated using Gaussian normal density, and GARCH_H is the GARCH model estimated using Hansen's skewed t. (2) The conditional VaR (CVaR) models are denoted in the form "Unconditional VaR-Variance Model", where Variance Model denotes model for σ_t^2 that is used to compute $z_t = \varepsilon_t/\sigma_t$. (3) MA = moving average; EWMA = exponentially weighted MA; HS = historic simulation; HWHS = Hull and White's (1998) modified HS; NPQ = nonparametric quantile regression of Cai (2002); MC = Monte Carlo; EVT = extreme value theory; GPD = generalized Pareto distribution; GEV = generalized extreme value distribution; GBM = geometric Brownian motion.

TABLE 2. Data

Country	Currency	In-Sample Period	R	Out-of-Sample Period	P
Thailand	Baht	January 1, 1992 - June 30, 1997	1434	July 1, 1997 - 6 15 1998	250
Korea	Won	December 1, 1992 - September 30, 1997	1262	October 1, 1997 - October 9, 1998	250
Russia	Rouble	May 7, 1993 - March 8, 1998	1326	March 9, 1998 - June 19, 1999	250
Turkey	Lira	Janaury 1, 1990 - March 31, 1994	1087	April 1, 1994 - December 14, 1995	250
Brazil	Real	January 1, 1993 - December 31, 1998	1106	January 1, 1999 - December 17, 1999	250

Notes: (1) Source: Data Stream. (2) Daily nominal foreign exchane rates.

TABLE 3. Christoffersen Test: Thailand and Korea

		Thailand						Korea					
		$\alpha = 0.05$			$\alpha = 0.01$			$\alpha = 0.05$			$\alpha = 0.01$		
k	Model	$\hat{p}_k^{0.05}$	LR_1	LR_2	$\hat{p}_k^{0.01}$	LR_1	LR_2	$\hat{p}_k^{0.05}$	LR_1	LR_2	$\hat{p}_k^{0.05}$	LR_1	LR_2
1	EWMA	.080	4.10	9.40	.024	3.58	2.423						
2	MA	.156	39.06	14.91	.088	58.43	14.794						
3	$GARCH_N$	.084	5.17	0.463	.044	15.96	1.017						
4	HS	.414	294.37	4.801	.165	159.04	15.116						
5	HS-$GARCH_N$	.008	14.04	0.032	.004	1.16	0.008						
6	MC	.309	170.98	14.812	.229	260.94	19.567						
7	MC-$GARCH_N$	.140	11.98	0.683	.064	33.26	0.001						
8	GPD	.892	1161.95	3.333	.189	195.71	28.824						
9	GPD-$GARCH_N$	.008	14.04	0.032	.004	1.164	0.008						
10	GEV10	.470	370.26	3.646	.430	648.109	3.718						
11	GEV10-$GARCH_N$	.008	14.04	0.032	.004	1.164	0.008						
12	GEV20	.470	370.26	3.646	.430	648.109	3.718						
13	GEV20-$GARCH_N$	.474	375.92	5.797	.004	1.164	0.008						
14	Hill	.205	73.39	17.810	.100	72.431	28.468						
15	Hill-$GARCH_N$	.040	0.54	0.837	.008	0.104	0.032						
16	HW	.470	370.26	11.052	.181	183.267	21.707						
17	Bootstrap-$GARCH_N$	.538	470.92	3.587	.052	22.404	1.433						
18	$CaViaR_S$	.153	36.57	13.232	.068	37.163	2.479						
19	$CaViaR_S$-$GARCH_N$	.165	44.23	21.885	.088	58.432	19.390						
20	$CaViaR_A$	.137	27.24	2.851	.080	49.592	1.186						
21	$CaViaR_A$-$GARCH_N$	.165	44.23	2.302	.080	49.592	1.186						
22	NPQ	.442	331.52	5.185	.189	195.709	12.482						
23	NPQ-$GARCH_N$	.008	14.04	0.032	.008	0.104	0.032						
24	$GARCH_H$	.092	7.61	13.300	.040	13.017	0.705						

TABLE 3. Christoffersen Test (continued): Russia and Turkey

		Russia						Turkey					
		$\alpha = 0.05$			$\alpha = 0.01$			$\alpha = 0.05$			$\alpha = 0.01$		
k	Model	$\hat{p}_k^{0.05}$	LR_1	LR_2	$\hat{p}_k^{0.01}$	LR_1	LR_2	$\hat{p}_k^{0.05}$	LR_1	LR_2	$\hat{p}_k^{0.05}$	LR_1	LR_2
1	EWMA							.052	0.03	14.02	.020	1.98	9.01
2	MA							.040	0.54	13.46	.012	0.10	5.43
3	GARCH_N							.068	1.58	8.46	.032	7.78	1.38
4	HS							.124	20.95	24.07	.048	19.09	5.70
5	HS-GARCH_N							.056	0.20	12.40	.012	0.10	0.07
6	MC							.080	4.10	17.28	.044	15.96	17.87
7	MC-GARCH_N							.060	0.52	10.95	.028	5.53	0.41
8	GPD							.996	1473.08	0.01	.068	37.16	38.24
9	GPD-GARCH_N							.161	41.61	56.86	.618	1089.23	8.29
10	GEV10							.149	34.14	13.93	.092	63.00	21.41
11	GEV10-GARCH_N							—-	–	–	–	–	–
12	GEV20							.169	46.90	25.10	.092	63.00	21.41
13	GEV20-GARCH_N							——	–	–	–	–	–
14	Hill							.104	11.98	1.38	.024	3.58	2.42
15	Hill-GARCH_N							.032	1.91	38.49	.012	0.10	0.07
16	HW							.149	34.14	22.27	.064	33.26	14.26
17	Bootstrap-GARCH_N							.221	86.48	22.27	.044	15.96	25.02
18	CaViaR_S							.165	44.23	11.55	.028	5.53	6.74
19	$\text{CaViaR}_S\text{-GARCH}_N$							.181	55.24	20.81	.076	45.33	19.11
20	CaViaR_A							.189	61.08	5.40	.060	29.50	10.95
21	$\text{CaViaR}_A\text{-GARCH}_N$							.241	103.89	2.11	.072	41.19	4.18
22	NPQ							.124	20.95	24.07	.060	29.50	6.72
23	NPQ-GARCH_N							.012	10.73	0.07	.056	25.88	12.40
24	GARCH_H												

TABLE 3. Christoffersen Test (continued):Brazil

		Brazil					
		$\alpha = 0.05$			$\alpha = 0.01$		
k	Model	$\hat{p}_k^{0.05}$	LR_1	LR_2	$\hat{p}_k^{0.01}$	LR_1	LR_2
1	EWMA	0.040	0.543	0.837	0.012	0.099	0.073
2	MA	0.052	0.025	9.575	0.044	15.961	12.555
3	GARCH_N	0.072	2.303	0.087	0.024	3.584	0.296
4	HS	0.281	141.947	8.196	0.060	29.499	16.874
5	HS-GARCH_N	0.072	2.303	0.087	0.012	0.099	0.073
6	MC	0.116	17.129	20.151	0.068	37.163	18.596
7	MC-GARCH_N	0.068	1.579	0.027	0.024	3.584	0.296
8	GPD	0.988	1441.731	0.073	0.052	22.404	14.875
9	GPD-GARCH_N	-	-	-	-	-	-
10	GEV10	0.426	310.096	4.165	0.277	345.211	4.571
11	GEV10-GARCH_N	0.004	18.402	0.008	0.004	1.164	0.008
12	GEV20	0.434	320.742	6.876	0.333	450.811	3.215
13	GEV20-GARCH_N	0.45	342.425	6.794	0.004	1.164	0.008
14	Hill	0.108	13.616	40.466	0.044	15.961	12.555
15	Hill-GARCH_N	0.032	1.907	0.531	0.004	1.164	0.008
15	HW	0.285	145.978	7.134	0.06	29.499	55.479
17	Bootstrap-GARCH_N	0.434	320.742	6.876	0.036	10.282	0.752
18	CaViaR_S	0.177	52.406	13.801	0.104	77.279	16.895
19	CaViaR_S-GARCH_N	0.177	52.406	16.786	0.133	113.481	13.759
20	CaViaR_A	0.177	52.406	8.664	0.092	63.004	17.592
21	CaViaR_A-GARCH_N	0.149	34.142	18.013	0.104	77.279	12.946
22	NPQ	0.321	183.993	3.290	0.068	37.163	18.596
23	NPQ-GARCH_N	0.012	10.734	0.073	0.080	49.592	0.306
24	GARCH_H	0.032	1.904	36.157	0.012	0.099	0.073

TABLE 4a. Reality Check: Thailand

k	Model	$A_k^{0.05}$	RC	$A_k^{0.01}$	RC	$B_k^{0.05}$	RC	$B_k^{0.01}$	RC	$C_k^{0.05}$	RC	$C_k^{0.01}$	RC	$D_k^{0.01}$	RC
1	EWMA	2.44	.052	4.09	.104	0.014	.325	0.020	.507	0.016	.964	0.014	.954		
2	MA	1.81	.127	2.93	.201	0.011	.472	0.013	.649	0.157	.548	0.235	.353		
3	GARCH$_N$	2.36	.036	3.90	.085	0.011	.438	0.013	.671	0.021	.958	0.064	.820		
4	HS	0.72	.388	1.97	.326	0.007	.498	0.010	.698	1.182	.000	0.639	.004		
5	HS+GARCH$_N$	16.11	.000	32.18	.000	0.024	.074	0.027	.188	0.056	.772	0.005	.908		
6	MC	1.01	.307	1.30	.467	0.008	.470	0.001	.755	0.687	.004	1.048	.000		
7	MC+GARCH$_N$	2.39	.034	3.49	.141	0.011	.420	0.013	.635	0.048	.913	0.134	.632		
8	GPD	0.234	1.000	2.11	.308	0.002	1.000	0.008	.884	4.666	.000	0.786	.000		
9	GPD-GARCH$_N$	10.61	.000	47.61	.000	0.024	.089	0.027	.175	0.056	.786	0.005	.930		
10	GEV10	0.58	.475	0.65	.990	0.006	.536	0.006	.967	1.487	.000	2.603	.000		
11	GEV10-GARCH$_N$	23.19	.000	40.92	.000	0.024	.091	0.027	.189	0.056	.778	0.005	.917		
12	GEV20	0.58	.458	0.64	1.000	0.006	.537	0.006	.974	1.487	.000	2.603	.000		
13	GEV20-GARCH$_N$	0.718	.442	41.20	.000	0.005	.612	0.027	.178	1.510	.000	0.005	.919		
14	Hill	1.50	.156	2.74	.200	0.009	.482	0.012	.679	0.295	.180	0.291	.220		
15	Hill-GARCH$_N$	4.07	.000	16.79	.000	0.014	.253	0.024	.232	0.002	.993	0.000	.980		
16	HW	0.70	.398	2.72	.249	0.006	.563	0.009	.815	1.487	.000	0.736	.001		
17	Bootstrap-GARCH$_N$	0.51	.503	3.55	.118	0.005	.574	0.012	.690	1.891	.000	0.090	.740		
18	CaViaR$_S$	1.09	.285	3.17	.189	0.008	.487	0.013	.639	0.507	.023	0.149	.582		
19	CaViaR$_S$-GARCH$_N$	1.05	.307	3.75	.133	0.008	.467	0.011	.673	0.538	.015	0.235	.350		
20	CaViaR$_A$	1.20	.226	3.08	.181	0.009	.486	0.013	.649	0.389	.079	0.199	.444		
21	CaViaR$_A$-GARCH$_N$	1.19	.235	2.98	.185	0.009	.500	0.012	.669	0.334	.144	0.199	.447		
22	NPQ	0.66	.432	1.81	.365	0.006	.521	0.010	.749	1.331	.000	0.786	.000		
23	NPQ-GARCH$_N$	23.62	.000	15.78	.000	0.024	.077	0.024	.232	0.056	.797	0.000	.989		
24	GARCH$_H$	1.06	.300	1.53	.409	0.006	.515	0.007	.969	0.961	.000	1.810	.000		

Notes: (1) The sample period of the data is from to December x, xxxx with the total xxxx observations. The models are estimated using $R = xxxx$ observations. (2) RC denotes the reality check p-value of White (2000) test computed using the stationary bootstrap of Politis and Romano (1994, PR). The Bootstrap Reality Check p-values are computed with 1,000 bootstrap resamples and the bootstrap smoothing parameter $q = 0.5$. See PR or White (1998) for the details. The p-values for $q = 0.25$ and 0.75 are similar and are not reported. RC is to compare all of the models except model k with model k ($k = 1, \ldots, 26$). That is to consider each model k as the benchmark model. (3) LR is the estimated statistics of Christoffersen (1998), distributed asymptotically $\chi(1)$ with the 5% critical value 3.84. Note that $LR = C_k^\alpha \times P$ and thus the loss $C_k^\alpha = P^{-1} \times LR$. (4) Loss D, D_k^α, is MRC, defined only for $\alpha = 0.01$ by the BIC.

TABLE 4b. Reality Check: Korea

k	Model	$A_k^{0.05}$	RC	$A_k^{0.01}$	RC	$B_k^{0.05}$	RC	$B_k^{0.01}$	RC	$C_k^{0.05}$	RC	$C_k^{0.01}$	RC	$D_k^{0.01}$	RC
1	EWMA														
2	MA														
3	GARCH$_N$														
4	HS														
5	HS+GARCH$_N$														
6	MC														
7	MC+GARCH$_N$														
8	GPD														
9	GPD-GARCH$_N$														
10	GEV10														
11	GEV10-GARCH$_N$														
12	GEV20														
13	GEV20-GARCH$_N$														
14	Hill														
15	Hill-GARCH$_N$														
16	HW														
17	Bootstrap-GARCH$_N$														
18	CaViaR$_S$														
19	CaViaR$_S$-GARCH$_N$														
20	CaViaR$_A$														
21	CaViaR$_A$-GARCH$_N$														
22	NPQ														
23	NPQ-GARCH$_N$														
24	GARCH$_H$														

Notes: See Table 4a.

TABLE 4c. Reality Check: Russia

k	Model	$A_k^{0.05}$	RC	$A_k^{0.01}$	RC	$B_k^{0.05}$	RC	$B_k^{0.01}$	RC	$C_k^{0.05}$	RC	$C_k^{0.01}$	RC	$D_k^{0.01}$	RC
1	EWMA														
2	MA														
3	GARCH$_N$														
4	HS														
5	HS+GARCH$_N$														
6	MC														
7	MC+GARCH$_N$														
8	GPD														
9	GPD-GARCH$_N$														
10	GEV10														
11	GEV10-GARCH$_N$														
12	GEV20														
13	GEV20-GARCH$_N$														
14	Hill														
15	Hill-GARCH$_N$														
16	HW														
17	Bootstrap-GARCH$_N$														
18	CaViaR$_S$														
19	CaViaR$_S$-GARCH$_N$														
20	CaViaR$_A$														
21	CaViaR$_A$-GARCH$_N$														
22	NPQ														
23	NPQ-GARCH$_N$														
24	GARCH$_H$														

Notes: See Table 4a.

TABLE 4d. Reality Check: Turkey

k	Model	$A_k^{0.05}$	RC	$A_k^{0.01}$	RC	$B_k^{0.05}$	RC	$B_k^{0.01}$	RC	$C_k^{0.05}$	RC	$C_k^{0.01}$	RC	$D_k^{0.01}$	RC
1	EWMA														
2	MA														
3	GARCH$_N$														
4	HS														
5	HS+GARCH$_N$														
6	MC														
7	MC+GARCH$_N$														
8	GPD														
9	GPD-GARCH$_N$														
10	GEV10														
11	GEV10-GARCH$_N$														
12	GEV20														
13	GEV20-GARCH$_N$														
14	Hill														
15	Hill-GARCH$_N$														
16	HW														
17	Bootstrap-GARCH$_N$														
18	CaViaR$_S$														
19	CaViaR$_S$-GARCH$_N$														
20	CaViaR$_A$														
21	CaViaR$_A$-GARCH$_N$														
22	NPQ														
23	NPQ-GARCH$_N$														
24	GARCH$_H$														

Notes: See Table 4a.

TABLE 4e. Reality Check: Brazil

k	Model	$A_k^{0.05}$	RC	$A_k^{0.01}$	RC	$B_k^{0.05}$	RC	$B_k^{0.01}$	RC	$C_k^{0.05}$	RC	$C_k^{0.01}$	RC	$D_k^{0.01}$	RC
1	EWMA	0.35	.000	3.44	.188	.0045	.058	0.029	.234	2.28	.000	0.000	.982		
2	MA	0.39	.000	3.73	.185	.0041	.877	0.021	.455	1.12	.724	0.064	.742		
3	GARCH_N	0.43	.000	2.69	.275	.0064	.000	0.022	.467	1.12	.906	0.014	.919		
4	HS	0.58	.000	2.59	.288	.0054	.000	0.018	.471	2.31	.000	0.118	.487		
5	HS+GARCH_N	0.35	.000	3.81	.128	.0055	.000	0.020	.502	1.12	.897	0.003	.985		
6	MC	0.33	.000	1.76	.394	.0042	.447	0.016	.490	3.81	.000	0.149	.369		
7	MC+GARCH_N	0.58	.000	2.51	.304	.0064	.000	0.022	.460	1.12	.889	0.014	.918		
8	GPD	0.00	.973	2.81	.276	-	-	0.019	.456	-	-	0.090	.625		
9	GPD-GARCH_N	-	-	-	-	-	-	-	-	-	-	-	-		
10	GEV10	0.39	.000	0.77	.859	.0041	.867	0.007	.889	1.12	.709	1.386	.000		
11	GEV10-GARCH_N	0.35	.000	41.58	.000	.0045	.061	0.034	.259	2.28	.000	0.005	.913		
12	GEV20	0.33	.000	0.66	1.000	.0042	.423	0.006	.999	3.81	.000	1.810	.000		
13	GEV20-GARCH_N	0.58	.000	40.23	.000	.0064	.000	0.034	.258	1.12	.904	0.005	.899		
14	Hill	0.43	.000	3.26	.231	.0064	.000	0.021	.457	1.12	.904	0.064	.731		
15	Hill-GARCH_N	0.58	.000	6.51	.004	.0054	.001	0.034	.269	2.31	.000	0.005	.912		
16	HW	0.40	.000	7.64	.022	.0041	.656	0.014	.551	1.10	.864	0.118	.510		
17	Bootstrap-GARCH_N	0.35	.000	2.59	.269	.0055	.000	0.010	.806	1.12	.905	0.041	.763		
18	CaViaR_S	0.44	.000	2.00	.358	.0054	.000	0.013	.591	2.31	.000	0.235	.183		
19	$\text{CaViaR}_S\text{-GARCH}_N$	0.39	.000	1.80	.392	.0041	.876	0.011	.679	1.12	.712	0..371	.050		
20	CaViaR_A	0.00	.963	2.09	.381	-	-	0.013	.600	-	-	0.235	.207		
21	$\text{CaViaR}_A\text{-GARCH}_N$	0.35	.000	2.04	.378	.0045	.070	0.013	.608	2.28	.000	0.235	.202		
22	NPQ	0.40	.000	1.80	.384	.0041	.650	0.016	.519	1.10	.897	0.149	.403		
23	NPQ-GARCH_N	0.44	.000	1.58	.423	.0054	.000	0.012	.655	2.31	.000	0.199	.259		
24	GARCH_H	0.33	.000	3.19	.215	.0004	.439	0.029	.225	3.81	.000	0.000	.985		

Notes: See Table 4a.