On Some Models for Value-at-Risk

Shusong Jin[1]

Email: jinss@fudan.edu.cn

Tel.: +86-21-65646570

[1]School of Management

Fudan University

670 Guosun Road, Shanghai 200433

P.R. China Wai Keung Li[2]

Email: hrntlwk@hku.hk

Tel.: +852-28592473

Philip L.H. Yu[2]

Email: plhyu@hku.hk

Tel.: +852-28578321

[2]Department of Statistics and Actuarial Science

The University of Hong Kong

Pokfulam Road, Hong Kong

This version: Nov. 1, 2006

Abstract

In recent years, value-at-risk (VaR) has become the standard tool for market risk measure and management. For better VaR estimation, Engle and Manganelli [2004] introduced the conditional autoregressive value-at-risk (CAViaR) model to estimate the VaR directly by quantile regression. To entertain the nonlinearity and structural change in the VaR, we extend the CAViaR idea using two approaches: the threshold GARCH (TGARCH) and the mixture-GARCH models. Some theoretical results are derived and methods for model estimation are proposed. Our models inherit all the advantages of the CAViaR model and enhance the nonlinear structure. The methods are applied to the S&P 500, Hang Seng, Nikkei and Nasdaq indices to illustrate our models.

Keywords: GARCH model, mixtures, threshold models, value-at-risk.

1 Introduction

Many financial crises happened without warning. These extreme price movements in the financial markets are rare, but they can bring fatal results to some corporations and disasters to a country's financial market. For instance, the New York stock market crashed in October 1987, and then, one decade later, the Asian stock market crashed also. Recent scandals of Enron had also caused the Dow Jones Industrial Average (DJIA) to drop sharply. These crises are the causes of ruin for hundreds of companies with their stock value evaporated in a short time. These markets and credit risk issues make investors more and more cautious. Nowadays, in the trend of globalization, market risk is of great importance and deserves more attention. Therefore, to a risk manager, a good measure of market risk is more than necessary. In particular, the Basel Committee on Banking Supervision [1996] at the Bank for International Settlements imposes on financial institutions such as banks and investment firms to meet capital requirements based on value-at-risk (VaR) estimates. As a result, value-at-risk has come out as a standard measurement of an institution's risk exposure and has become a standard tool to evaluate and manage financial risk.

Value-at-risk is mainly concerned with market risk. It can be defined as the maximal loss of a position on a financial asset during a given time period for a given probability. Let $\Delta V(l)$ be the change in value in the position from time t to $t + l$. We can define the VaR of the position over the time horizon l with

probability η as

$$\eta = P[\Delta V(l) \leqslant \text{VaR}] \equiv F_l(\text{VaR}), \tag{1}$$

where $F_l(x)$ is the cumulative distribution function (CDF) of $\Delta V(l)$, evaluated at $\Delta V(l) = x$. See Jorion [2001] for a comprehensive review of VaR. Since the potential loss of holding a financial asset can be represented by its return, we can re-define the VaR in terms of the return. Let $x_{t+l} = \log P_{t+l} - \log P_t$ be the logarithm return of holding the asset from time t to $t + l$, where P_t is the asset price at time t. Then the VaR over the time horizontal with probability η is given by

$$\eta = P[x_{t+l} \leqslant \text{VaR}] \equiv F_l(\text{VaR}), \tag{2}$$

$F_l(x)$ is the CDF of x_{t+l}, evaluated at $x_{t+l} = x$. Although the concept of VaR is not difficult to understand, the estimation of VaR is a very challenging statistical problem. The pioneer is the risk management group at J. P. Morgan who developed the RiskMetrics model for measuring VaR in 1995. The RiskMetrics model assumes that returns of a financial asset follow a conditional normal distribution with zero mean and variance being expressed as an exponentially weighted moving average of historical squared returns. Because of the many nice properties of the normal distribution, the calculation of VaR is relatively straightforward and it can also be applied to a long horizon VaR estimation (Jorion [2001], Wong and So [2003]). However, as many papers have pointed out, the RiskMetrics model fails to measure market risk accurately because this model lack nonlinear property which is a significant feature of the financial market. To overcome

this shortcoming, various methods, such as that based on GARCH (Bollerslev [1986]) and its variants such as integrated GARCH (Engle and Bollerslev [1986]) and exponential GARCH (Nelson [1991]) and extreme value theories (Danielsson and de Vries [1997]) were proposed and found to have significant improvements. Manganelli and Engle [2001] summarized some recent development of the VaR estimation.

Generally speaking, from a statistical point of view, VaR defined by (2) is just the conditional η-th quantile of the distribution of a future asset return. This feature enlightens us to estimate the quantile directly. Koenker and Bassett [1978] developed the quantile regression method to solve this question. Their method provides a nonparametric approach to VaR estimation. The advantage of this method is that it makes no specific distributional assumption on the return of the asset.

For a given value η, $0 \leqslant \eta \leqslant 1$, the η-th quantile of the a variable x is defined as $C_\eta(x) = \inf\{x|F(x) \geqslant \eta\}$, where F is the CDF of x. Koenker and Bassett [1978] suggested that the quantile of $b = C_\eta(x)$ based on a sample of *iid* realization of x $\{x_t\}$ can be estimated by solving the following minimization problem

$$\min_{b\in\mathscr{R}} \left[\sum_{t\in\{t:x_t\geqslant b\}} \eta|x_t - b| + \sum_{t\in\{t:x_t<b\}} (1-\eta)|x_t - b| \right]. \tag{3}$$

If the returns series $\{x_t\}$ follows an autoregressive process of order k:

$$x_t = \phi_0 + \sum_{i=1}^{k} \phi_i x_{t-i} + u_t \tag{4}$$

the $100\eta\%$ conditional VaR of x_t is then given by

$$\mathrm{VaR}_t(\eta) = \phi_0 + \sum_{i=1}^{k} \phi_i x_{t-i} + C_\eta(u_t|\mathscr{F}_{t-1}), \tag{5}$$

where $C_\eta(u_t|\mathscr{F}_{t-1})$ is the η-th conditional quantile of the residual process u_t, and $\mathscr{F}_{t-1}$ is the information up to time $t-1$. With the quantile regression method, Wu and Xiao [2002] obtained VaR process for the AR model and suggested this method can be applied to ARCH model.

Engle and Manganelli [2004] applied this method to the VaR problem and proposed a conditional value-at-risk (CAViaR) model with a general specification for VaR at time t

$$\mathrm{VaR}_t = f(x_t, \beta_\theta) = \beta_0 + \sum_{i=1}^{p} \beta_i \,\mathrm{VaR}_{t-i} + l(\beta_{p+1}, \ldots, \beta_{p+q}; \mathscr{F}_{t-1}) \tag{6}$$

where $l(\cdot)$ is a pre-specified function and θ is the subscript suppressed for notational convenience. They suggested the first order CAViaR model:

$$\mathrm{VaR}_t = \beta_0 + \beta_1 \,\mathrm{VaR}_{t-1} + l(\beta_2, x_{t-1}, \mathrm{VaR}_{t-1}) \tag{7}$$

is sufficient for practical use. Some examples are listed below:

1. Adaptive: $\mathrm{VaR}_t = \mathrm{VaR}_{t-1} + \beta[I(x_{t-1} \leqslant -\mathrm{VaR}_{t-1}) - \theta]$.

2. Symmetric absolute value: $\mathrm{VaR}_t = \beta_0 + \beta_1 \,\mathrm{VaR}_{t-1} + \beta|x_{t-1}|$.

3. Direct GARCH(1,1)[1]: $\mathrm{VaR}_t = (\beta_0 + \beta_1 \,\mathrm{VaR}_{t-1}^2 + \beta_2 x_{t-1}^2)^{1/2}$.

[1]Engle and Manganelli [2004] named this model indirect GARCH(1,1) model. However, in this paper, we would like to call it direct GARCH(1,1) model because it models the VaR_t directly.

As return series may usually exhibit nonlinearity such as a threshold structures or mixture models, this paper aims at extending CAViaR models to incorporate nonlinear structure of return series in order to improve the VaR estimation.

In the literature, examples of nonlinearity abound. For instance, DeBondt and Thaler [1985] pointed out that a portfolio of "loser" stocks may outperform a portfolio of "winner" stocks. Their article shown that a simple trading strategy of buying "recent losers" and selling "recent winners" could yield a substantial excess profit in both short-term and long-term investment horizons. That means stock returns always exhibit an apparent asymmetric reverting pattern in their return behavior. French et al. [1987] found evidence that the expected market risk premium is positively related to the predictable volatility of stock returns, but the unexpected stock market returns are negatively related to the unexpected change in the volatility of stock returns. Nam et al. [2002] argued that the relative profitability of "loser" stocks is attributed to an asymmetric reverting property of their return dynamics. Chan and Maheu [2002] used conditional jump dynamics to explain the behavior of the stock market returns. From these results, it therefore appears natural to assume that the underlying data has an asymmetric structure. It is well known that threshold or mixture models readily allow for asymmetric structures. Furthermore, changing volatility (Schwert and Stambaugh [1987]) is also an important stylized fact of stock returns. Hence models with nonlinearity and changing volatility structures are needed to be considered in risk management.

Two popular models for describing the changing volatilities are the autoregressive conditional heteroscedastic (ARCH) model (introduce by Engle [1982]) and the generalized autoregressive conditional heteroscedastic (GARCH) model (Bollerslev [1986]). To incorporate nonlinear structure, many of their variants were proposed in the literature including models for asymmetric behavior of variance such as the EGARCH (Nelson [1991]) and TARCH (Rabemananjara and Zakoïan [1993] and Zakoïan [1994]). Li and Li [1996] suggested the double-threshold ARCH (DTARCH) model which is an extension of Tong and Lim [1980]. The model can handle the situation where both the conditional mean and the conditional variance specifications are piecewise linear. Another variant is the class of mixture models recently proposed by Wong and Li [2000, 2001] in which the return is assumed to follow a mixture of k components AR-ARCH model. The use of "mixtures" model is not new and has been considered in existing VaR literatures (for instance, Hull [2002]), but their models do not allow changing volatility. In this paper, we will focus on the conditional VaR estimation by extending CAViaR (Engle and Manganelli [2004]) models to include the threshold GARCH (TGARCH) (an extension of Li and Li's double TARCH, 1996) and mixture-GARCH (an extension of Wong and Li [2001]'s mixture-ARCH) models since these two models can better explain the asymmetric phenomenon in the financial market. We call the new models direct-VaR models. The organization of this paper is as follows. Section 2 reviews TGARCH and mixture-GARCH models. Section 3 formulates two direct-VaR models based on either TGARCH or mixture-GARCH models. The estimation methods of the proposed direct-VaR

models are introduced in Section 4. Simulations are done in Section 5 to examine the effectiveness of our proposed estimation methods. The proposed models are applied in Section 6 to the S&P 500 index to illustrate the capability of the models in estimating VaR. For comparison, we also apply our models to Hang Seng, Nikkei and Nasdaq indices which can be regarded as immature markets and indices. Section 7 concludes the advantage of our models.

2 The TGARCH and Mixture-GARCH model

We first consider threshold type models for a time series x_t. Tong and Lim [1980] first suggested the self-exciting threshold autoregressive model (SETAR) model for nonlinear time series dynamic structure. A time series x_t is said to follow a k-regime SETAR model with threshold variable x_{t-d} if it satisfies

$$x_t = \phi_0^{(j)} + \phi_1^{(j)} x_{t-1} + \cdots + \phi_p^{(j)} x_{t-p} + a_t^{(j)}, \quad \text{if } \gamma_{j-1} \leqslant x_{t-d} < \gamma_j,\ j = 1, \ldots, k, \qquad (8)$$

where k and d are positive integers, $j = 1, \ldots, k$. The thresholds γ_i are real numbers such that $-\infty = \gamma_0 < \gamma_1 < \cdots < \gamma_{k-1} < \gamma_k = \infty$, where the superscript (j) is used to signify the regime. The series $\{a_t^{(j)}\}$ are *iid* error with mean 0 and variance σ_j^2 and are mutually independent for different j. The parameter d is referred to as the delay parameter and γ_js as the thresholds. With the similar idea, to capture the asymmetric volatility characteristic of financial time series, we extend the double threshold ARCH (DTARCH) of Li and Li [1996] to a general

k-regime threshold GARCH (TGARCH) model as below:

$$x_t = \varepsilon_t \sqrt{h_t}, \tag{9a}$$

$$h_t = \alpha_0^{(j)} + \sum_{r=1}^{p_j} \alpha_r^{(j)} x_{t-r}^2 + \sum_{i=1}^{q_j} \beta_i^{(j)} h_{t-i}, \qquad \gamma_{j-1} \leqslant x_{t-d} < \gamma_j, j = 1, \ldots, k, \tag{9b}$$

where regime number k and delay parameter d are positive integers, $j = 1, \ldots, k$. γ_j's are real numbers such that $-\infty = \gamma_0 < \gamma_1 < \cdots < \gamma_{k-1} < \gamma_k = \infty$, the superscript (j) is used to signify the regime, $\{\varepsilon_t\}$ are *iid* noise with mean 0 and unit variance. We denote model (9) as TGARCH(p_1, p_2, ..., p_k; q_1, q_2, ..., q_k). With some regularity conditions, Liu et al. [1997] proved that model (9) is stationary and ergodic. One advantage of the TGARCH model is that it can model asymmetric and limit cycle behavior in a natural way and break the homoscedastic variance restriction as ordinary SETAR-type models do.

An alternative nonlinear time series model is the mixture AR-ARCH model proposed by Wong and Li [2001]. They assume that x_t follows a mixture of K-component models with each components being a AR-ARCH model. The general K-component model is given by:

$$F(x_t|\mathscr{F}_{t-1}) = \sum_{k=1}^{K} \pi_k \Phi\left(\frac{e_{k,t}}{\sqrt{h_{k,t}}}\right) \tag{10a}$$

$$e_{k,t} = x_t - \phi_{k0} - \phi_{k1} x_{t-1} - \cdots - \phi_{kp_k} x_{t-p_k} \tag{10b}$$

$$h_{k,t} = \beta_{k0} + \beta_{k1} e_{k,t-1}^2 + \cdots + \beta_{kq_k} e_{k,t-q_k}^2, \tag{10c}$$

where $\mathscr{F}_{t-1}$ is the information set up to time $t-1$; $\Phi(\cdot)$ is the (conditional) cumulative distribution function of the standard Gaussian distribution; and $\pi_1 + \cdots +$

$\pi_K = 1$, $\pi_k > 0$ $(k = 1, \ldots, K)$. Here $F(x_t|\mathscr{F}_{t-1})$ is the conditional cumulative distribution function of x_t given the past information. They denote this model as the MAR-ARCH(K; p_1, $p_2, \ldots, p_K$; q_1, $q_2, \ldots, q_K$). We extend the ARCH formulation to GARCH and the mixture-GARCH model with K components is defined as

$$x_t = \varepsilon_t \sqrt{h_t} \tag{11a}$$

$$h_t = \sum_{k=1}^{K} Z_{kt} \left(\alpha_{k0} + \sum_{i_k=1}^{q_k} \alpha_{ki_k} x_{t-i_k}^2 + \sum_{j_k=1}^{p_k} \beta_{kj_k} h_{t-j_k} \right), \tag{11b}$$

where $Z_{kt} = 1$ when the x_t is drawn from the k-th component (let π_k be the probability that $Z_{kt} = 1$), $Z_{kt} = 0$ otherwise. Then $\pi_1 + \pi_2 + \cdots + \pi_K = 1$, $\pi_k > 0$, $(k = 1, \ldots, K)$; ε_t is *iid* white noise following the standard normal distribution. We call model (11) a mixture-GARCH(K; p_1, $p_2, \ldots, p_K$; q_1, $q_2, \ldots, q_K$). As in the case for the GARCH(1,1) model, the most important special case of the mixture-GARCH model is the (2;1,1;1,1) model,

$$x_t = \varepsilon_t \sqrt{h_t}, \tag{12a}$$

$$\begin{aligned} h_t = {} & Z_t(\alpha_{10} + \alpha_{11} x_{t-1}^2 + \beta_{11} h_{t-1}) \\ & + (1 - Z_t)(\alpha_{20} + \alpha_{21} x_{t-1}^2 + \beta_{21} h_{t-1}) \end{aligned} \tag{12b}$$

where Z_t follows the Bernoulli distribution:

$$Z_t = \begin{cases} 1 & \text{with probability } \pi, \\ 0 & \text{with probability } (1 - \pi). \end{cases} \tag{12c}$$

We denote the model (12) as the mixture-GARCH(1,1) with two components.

3 Direct-VaR Models

Following the idea of quantile regression, Engle and Manganelli [2004] proposed to model directly the quantiles instead of the underlying volatility. A conditional autoregressive quantile regression model, termed as the conditional autoregressive value-at-risk (CAViaR) model, is introduced. The general form of this model is

$$\mathrm{VaR}_t = f(x_t, \beta_\theta) = \beta_0 + \sum_{i=1}^{p} \beta_i \,\mathrm{VaR}_{t-1} + l(\beta_{p+1}, \ldots, \beta_{p+q}; \mathscr{F}_{t-1}), \tag{13}$$

By analogy with the popular GARCH(1,1) model, the GARCH(1,1) direct-VaR model (Engle and Manganelli, 2004, section 3) is defined as

$$\mathrm{VaR}_t = (\beta_0 + \beta_1 \,\mathrm{VaR}_{t-1}^2 + \beta_2 x_{t-1}^2)^{1/2}. \tag{14}$$

Engle and Manganelli [2004] pointed out that if the underlying volatilities were truly generated from a GARCH(1,1) process, the VaR follows the GARCH(1,1) process (14). Note that (14) is not a GARCH(1,1) model in the usual sense. It is in fact a "direct" model for VaR_t. We will call (14) a GARCH(1,1) direct-VaR model. In the following, we obtain a similar result for direct-VaR models when the underlying data are truly a TGARCH or mixture-GARCH(1,1) model respectively.

Theorem 3.1. *If x_t follows a TGARCH process* (9)*, then the* VaR_t^2 *process follows a TGARCH process* (15)*. If x_t follows a mixture-GARCH(1,1) process* (12)*, then the VaR process also follows an mixture-GARCH(1,1) process* (16)*.*

$$\mathrm{VaR}_t^2 = a_0^{(j)} + \sum_{r=1}^{p_j} a_r^{(j)} x_{t-r}^2 + \sum_{i=1}^{q_j} b_i^{(j)} \,\mathrm{VaR}_{t-i}^2 \quad \gamma_{j-1} \leqslant x_{t-d} < \gamma_j, \tag{15}$$

$$\mathrm{VaR}_t^2 = Z_t\big(a_{10} + a_{11}x_{t-1}^2 + b_{11}\,\mathrm{VaR}_{t-1}^2\big)$$
$$+(1 - Z_t)\big(a_{20} + a_{21}x_{t-1}^2 + b_{21}\,\mathrm{VaR}_{t-1}^2\big), \tag{16}$$

where C_η is the lower (left-tail) η-th quantile of the standard normal distribution, $a_r^{(j)} = C_\eta^2\alpha_r^{(j)}$, $r = 0, \ldots, p_j$, $b_i^{(j)} = \beta_i^{(j)}$, $i = 1, \ldots, q_j$, $j = 1, \ldots, k$; Z_t follows (12c)*; $a_{k0} = C_\eta^2 \cdot \alpha_{k0}$, $a_{k1} = C_\eta^2 \cdot \alpha_{k1}$ and $b_{k1} = \beta_{k1}$, $k = 1, 2$.*

Corollary 3.2. *If under the assumptions of theorem 3.1, the η-th* VaR_t *follows* (15) *or* (16)*, then the μ-th* VaR_t *follows* (17) *or* (18) *respectively.*

$$\mathrm{VaR}_t^2 = m^2 \cdot a_0^{(j)} + \sum_{r=1}^{p_j} m^2 \cdot a_r^{(j)} x_{t-r}^2 + \sum_{i=1}^{q_j} b_i^{(j)}\,\mathrm{VaR}_{t-i}^2 \qquad \gamma_{j-1} \leqslant x_{t-d} < \gamma_j, \tag{17}$$

$$\mathrm{VaR}_t^2 = Z_t(m^2 \cdot a_{10} + m^2 \cdot a_{11}x_{t-1}^2 + b_{11}\,\mathrm{VaR}_{t-1}^2)$$
$$+ Z_t^c(m^2 \cdot a_{20} + m^2 \cdot a_{21}x_{t-1}^2 + b_{21}\,\mathrm{VaR}_{t-1}^2), \tag{18}$$

where $m = C_\mu/C_\eta$, $Z_t^c = 1 - Z_t$ and $j = 1, \ldots, k$.

Remark. Corollary 3.2 is very useful because if we have found the $100\eta\%$ VaR process and want to know the $100\mu\%$ VaR process for $\mu \neq \eta$, we need not fit the model again. We can apply corollary 3.2 above and obtain the result directly.

4 Model Estimation for Threshold and Mixture Type Direct-VaR Models

4.1 The Direct-VaR TGARCH Model

There are many ways to estimate a threshold model. One of the most popular methods proposed by Tsay [1989] is to estimate the parameters by the structural change method. However, in our case, it is hard to apply. Although the data may actually have a structural change, an accurate numerical procedure is difficult to implement. We resort to use a more direct but slightly more time consuming method for the TGARCH model estimation. Firstly, we sort the data points into an ascending order. Then we remove the smallest 25% and the largest 25% data pionts and consider each data point in the middle 50% as possible candidate value for the threshold parameter. As an example, we illustrate the method to fit a direct-VaR TGARCH(1,1;1,1) model for the underlying series x_t with two regimes. Following (15), let

$$\mathrm{VaR}_t^2 = \begin{cases} a_0^{(1)} + a_1^{(1)} x_{t-1}^2 + b_1^{(1)} \mathrm{VaR}_{t-1}^2, & x_{t-d} < \gamma, \\ a_0^{(2)} + a_1^{(2)} x_{t-1}^2 + b_1^{(2)} \mathrm{VaR}_{t-1}^2, & x_{t-d} \geqslant \gamma. \end{cases} \tag{19}$$

The initial value VaR_1 is determined by C_η, the η-th quantile of the standard normal distribution, times the sample standard error. Other VaR_t can be determined iteratively by (19). If $C_\eta < 0$, we take the negative square root. If $C_\eta > 0$, we choose the positive square root. For a given threshold candidate γ, we obtain

the conditional η-th quantile by minimizing

$$L = \left\{ \frac{\left\{ \sum_{x_t \geqslant \mathrm{VaR}_t} \eta |x_t - \mathrm{VaR}_t| + \sum_{x_t < \mathrm{VaR}_t} (1-\eta) |x_t - \mathrm{VaR}_t| \right\} \cdot I(x_{t-d} < \gamma)}{\sum_{t=d+1}^{n} I(x_{t-d} < \gamma)} \right\}$$
$$+ \left\{ \frac{\left\{ \sum_{x_t \geqslant \mathrm{VaR}_t} \eta |x_t - \mathrm{VaR}_t| + \sum_{x_t < \mathrm{VaR}_t} (1-\eta) |x_t - \mathrm{VaR}_t| \right\} \cdot I(x_{t-d} \geqslant \gamma)}{\sum_{t=d+1}^{n} I(x_{t-d} \geqslant \gamma)} \right\},$$

where $I(\cdot)$ is the indicator function. The first term can be interpreted as the objective function for the conditional quantile function corresponding to the regime $x_{t-d} < \gamma$ normalized by the number of observations in that regime. A similar interpretation holds for the second term which corresponds to the regime $x_{t-d} \geqslant \gamma$. The parameter d is chosen from natural numbers. Since in practice, the d seldom exceeds 2, so we choose 1, 2 and 3 as its candidates. We repeat the method above for all the candidates and choose the threshold and parameter values that minimize L globally.

4.2 The Mixture-GARCH(1,1) Direct-VaR Model

Motivated by the Expectation-Maximization (EM) algorithm, we suggest a two-stage estimation to find the η-th quantile for VaR estimation. Let

$$L = \left\{ \sum_{x_t \geqslant \mathrm{VaR}_t} \eta \, |x_t - \mathrm{VaR}_t| + \sum_{x_t < \mathrm{VaR}_t} (1-\eta) \, |x_t - \mathrm{VaR}_t| \right\}. \tag{20}$$

As in White [1994], here we treat L as a generalized objective function such that the minimization of L is equivalent to the approach such that an EM style algorithm can be applied to obtain the quasi maximum likelihood estimates (QMLE).

We assume that x_t also follows a mixture-GARCH(1,1) model corresponding to theorem 3.1. For a mixture-GARCH(1,1) model (16), it is clear that the Z_ts are unobserved and hence can be regarded as missing information. Therefore, the two-stage algorithm can be used to handle this type of parameter estimation. We divide all the parameters into two groups, the parameters containing missing information and the other parameters not containing it. We call them group 1 and group 2 respectively. After that, our algorithm contains two stages. In stage 1, we suppose that all the parameters in group 2 are given by the previous step, and use them as initial values to estimate the missing information parameters in group 1 by taking conditional expectations of these parameters in group 1 with respect to the values of parameters in group 2. In stage 2, we use the conditional expectation of parameters under the mixture-GARCH model for x_t in group 1 as the real value of the missing information and estimate the parameters in group 2 by maximizing the likelihood function. These two steps are repeated until all the parameters have converged. Here, the parameter π corresponds to the missing information, so it is in group 1. The other parameters a_{k0}, a_{k1} and b_{k1}, $k = 1, 2$ are in group 2.

Stage 1. Given parameter estimates $\hat{a}_{k0}$, $\hat{a}_{k1}$, $\hat{b}_{k1}$, $k = 1, 2$, from the previous step, we estimate $\hat{\pi}_k$ as follows. We denote T_{kt} as the conditional probability of x_t in the k-th component, $k = 1, 2$. Since $\sqrt{h_t} = \mathrm{VaR}_t / C_\eta$, following Wong and Li [2001, section 3], we obtain

$$T_{kt} = \frac{\pi_k}{\mathrm{VaR}_t^{(k)}} \phi\left(\frac{C_\eta \cdot x_t}{\mathrm{VaR}_t^{(k)}}\right) \left(\sum_{l=1}^{2} \frac{\pi_l}{\mathrm{VaR}_t^{(l)}} \phi\left(\frac{C_\eta \cdot x_t}{\mathrm{VaR}_t^{(l)}}\right)\right)^{-1},$$

where $\phi(\cdot)$ is the density of the standard normal distribution,

$$\mathrm{VaR}_t^{(k)} = \mathrm{Sgn}(C_\eta)(\hat{a}_{k0} + \hat{a}_{k1}x_{t-1}^2 + \hat{b}_{k1}\mathrm{VaR}_{t-1}^2)^{1/2},$$

and $\mathrm{Sgn}(x)$ is 1, 0, -1 if $x > 0$, $x = 0$ and $x < 0$ respectively. T_{kt} can be computed by using the estimates of π_1 and π_2 obtained from the previous iteration. Therefore, we can estimate π_k, $k = 1, 2$ by

$$\hat{\pi}_k = \frac{\sum_{t=3}^{n} T_{kt}}{n-2}, \qquad k = 1, 2.$$

Stage 2. Now, using $\hat{\pi}_1$ and $\hat{\pi}_2$ obtained in Stage 1, a_{k0}, a_{k1}, b_{k1}, $k = 1, 2$ are updated by minimizing L where

$$L = \left\{ \sum_{x_t \geqslant \widehat{\mathrm{VaR}}_t} \eta \left| x_t - \widehat{\mathrm{VaR}}_t \right| + \sum_{x_t < \widehat{\mathrm{VaR}}_t} (1-\eta) \left| x_t - \widehat{\mathrm{VaR}}_t \right| \right\},$$

and

$$\begin{aligned} \widehat{\mathrm{VaR}}_t = &\mathrm{Sgn}(C_\eta)\big(\hat{\pi}_1(a_{10} + a_{11}x_{t-1}^2 + b_{11}\widehat{\mathrm{VaR}}_{t-1}^2)^{1/2} \\ &+ \hat{\pi}_2(a_{20} + a_{21}x_{t-1}^2 + b_{21}\widehat{\mathrm{VaR}}_{t-1}^2)^{1/2}\big). \end{aligned} \tag{21}$$

Since we can not determine the source component of x_t, $t = 1, \ldots, n$, we have to use the expectation of Z_t and $1 - Z_t$, $\hat{\pi}_1$ and $\hat{\pi}_2$, to estimate VaR_t. Therefore $\hat{\mathrm{VaR}}_t$ is in fact the conditional expectation of the true VaR_t. We repeat these two steps until the parameter estimates converge.

Remark. In probability theory, the parameters of model (16) should be estimated based on the information $\mathscr{F}_{t-1}$ which is the σ-field generated by observations $(x_1, \ldots, x_{t-1})$. In our estimation method, at the first glance, the parameters seem to be estimated based on the information $(x_1, \ldots, x_{t-1}, z_1, \ldots, z_{t-1})$. However, the

conditional expectation of the missing information $z_1, \ldots, z_{t-1}$ is replaced by the conditional expectation and can be treated as a function of $x_1, \ldots, x_{t-1}$ from the method of our estimation. Therefore, our method can be regarded as a generalized estimation method based on $\mathscr{F}_{t-1}$.

From the discussion above, it can be proved that under some regular conditions (White [1994, page 75] or Jin [2005, page 63]), the estimators we proposed are consistent under the probability sense. Furthermore, we find that no assumption on the distribution of the error terms is needed, hence reducing the risk of misspecification. Moreover, as Manganelli and Engle [2001] pointed out, even if the quantile process is misspecified, the minimization of the regression quantile objective function can still be interpreted as the minimization of the Kullback-Leibler information criterion, which measures the discrepancy between the true model and the one under study. Our models inherit all these advantages and it is expected that after incorporating a flexible nonlinear structure on VaR, the VaR can be estimated more accurately.

5 Simulations

In this section we simulate two time series data sets from the following two models and fit the direct-VaR$_t$ process with TGARCH and mixture-GARCH model by quantile regression. The models are

Model 1.

$$
\begin{aligned}
x_t &= \varepsilon_t \sqrt{h_t}, \\
h_t &= \begin{cases} 0.2 + 0.25 x_{t-1}^2 + 0.3 h_{t-1} & x_{t-1} < 0, \\ 0.1 + 0.15 x_{t-1}^2 + 0.4 h_{t-1} & x_{t-1} \geqslant 0. \end{cases}
\end{aligned} \tag{22}
$$

Model 2.

$$
\begin{aligned}
x_t &= \varepsilon_t \sqrt{h_t}, \\
h_t &= Z_t (0.1 + 0.1 x_{t-1}^2 + 0.4 h_{t-1})^{1/2} \\
&\quad + (1 - Z_t)(0.5 + 0.2 x_{t-1}^2 + 0.3 h_{t-1})^{1/2},
\end{aligned} \tag{23}
$$

where $Z_t = 1$ with probability 0.3 and $Z_t = 0$ with probability 0.7.

Model 1 is a TGARCH(1,1;1,1) model with two regimes and Model 2 is a two-component mixture-GARCH(1,1) models. For each model, we generated a white noise $\{\varepsilon_t\}$, $t = 1, \ldots, 1001$, and simulated a time series realization of length 1001 using (22) or (23). Let the series be $\{x_t\}$, $t = 1, \ldots, 1001$. The $\{x_t\}$, $t = 1, \ldots, 1000$ were used for model fitting. The last one, x_{1001}, was used to test the accuracy of the model. Initial values for estimation were drawn uniformly from a small interval containing the true parameters in the center. By repeating the above procedure 1500 times, we obtained 1500 one-step VaR's for testing the accuracy of the VaR prediction. The results are listed in Table 1. The upper section contains the result for TGARCH(1,1;1,1) model (22). We found that the means of the estimated parameters were very close to the true value. The number of times x_{1001} smaller than the estimated 5% VaR is 72. That is about 4.8% of

the times that the 5% quantile value was exceeded. This is very close to the theoretical value of 5%.

The lower part of Table 1 contains the results for mixture-GARCH model. Although the estimated parameters for the VaR process were slightly different from the true value, only 76 x_{1001} are outside of the estimated VaR. That is about 5.1% of the 5% quantile value was exceeded. From Table 1, we can find that the quantile regression method for the direct-VaR estimation in TGARCH and mixture-GARCH(1,1) cases can provide accurate VaR estimates.

6 Empirical Results

In this section, we applied our model to four real data sets: S&P 500 daily closing indices from Jan. 2, 1998 to Aug. 31, 2004, Hang Seng daily closing indices of Hong Kong from Jan. 2, 1997 to Dec. 31, 2003, Nasdaq daily closing indices from Jan. 2, 1997 to Dec. 31, 2003 and Nikkei daily closing indices from Jan. 6, 1998 to Dec. 30, 2003. We will discuss the study of S&P 500 in details and other three examples will be explained briefly. All these data sets are obtained from Yahoo.

The first example is the daily S&P 500 closing indices from Jan. 2, 1998 to Aug. 31, 2004. Days that the market was closed were removed and there are totally 1674 observations in the series. We denoted them as $\{y_t\}$ and used these observations for our empirical research.

Let the daily log-returns of the S&P 500 $\{r_t\}$ be defined by $r_t = \log y_t - \log y_{t-1}$. With the log-transformation, all $\{r_t\}$ are around zero and stationary. Figure 1 shows the S&P 500 series and its return series respectively. The autocorrelation (ACF) and partial autocorrelation (PACF)[2] functions of the return series $\{r_t\}$ do not show any apparent autoregressive structure and therefore, we will use the return series $\{r_t\}$ to fit the model directly without any linear filter. To illustrate the capability of our method, we choose the first 1173 log-returns, r_1, $\ldots$, r_{1173}, for model fitting. The remaining 500 data points, r_{1174}, $\ldots$, r_{1673} are used to check the number of data points outside the predicted VaR_t.

The first direct-VaR model we fit to the $r_1, \ldots, r_{1173}$ was the TGARCH(1,1;1,1) direct-VaR. The method was the same as what had been described in subsection 4.1. As mentioned before, we sorted the log-return in ascending order, discard the smallest and largest 25% and tried every r_t in the middle 50% as candidates for the threshold and obtained a TGARCH(1,1;1,1) model (24). For the sake of checking accuracy of long position VaR estimation, we did not update the model though new data points $r_{1174}, \ldots, r_{1673}$ came into the system daily during the checking stage. The second direct-VaR model considered was the mixture-GARCH(1,1) direct-VaR model (25). We applied the two-stage quantile regression method which had been described in subsection 4.2 to fit the mixture-GARCH(1,1) direct-VaR model. We summarized the result as below:

[2]The plots of the ACF and PACF can be found in Jin [2005]

TGARCH(1,1;1,1) direct-VaR:

$$\mathrm{VaR}_t^2 = \begin{cases} 2.413 \times 10^{-6} + 7.938 \times 10^{-2} r_{t-1}^2 & \\ \quad +9.072 \times 10^{-1}\,\mathrm{VaR}_{t-1}^2 & r_{t-1} \leqslant -1.343 \times 10^{-2}, \\ 4.176 \times 10^{-4} + 2.135 \times 10^{-2} r_{t-1}^2 & \\ \quad +1.472 \times 10^{-1}\,\mathrm{VaR}_{t-1}^2 & r_{t-1} > -1.343 \times 10^{-2}. \end{cases} \tag{24}$$

Since this is the left tail VaR, we choose the negative square roots. There are 31 return values lying outside of the 5% VaR. This amounts to an empirical coverage of 6.2%.

Mixture-GARCH(1,1) direct-VaR:

$$\begin{aligned} \mathrm{VaR}_t^2 = {} & z_t(6.2713 \times 10^{-4} + 0.1896 r_{t-1}^2 + 0.05369\,\mathrm{VaR}_{t-1}^2) \\ & + (1 - z_t)(6.6137 \times 10^{-4} + 0.25093 r_{t-1}^2 + 0.6733\,\mathrm{VaR}_{t-1}^2) \end{aligned} \tag{25}$$

where $z_t = 1$ with probability 0.1435 and $z_t = 0$ with probability 0.8565.We also choose the negative square root for VaR_t. There are 28 return values lying outside of the 5% VaR. This amounts to an empirical coverage of 5.6%.

Since S&P 500 is a developed indices, some performance checking of our model in immature markets and indices are also needed. We choose three of these indices, Hang Seng index (Figure 2) of Hong Kong, Nikkei indices of Japan and Nasdaq indices for technological stocks. The periods for these three indices are seven years, from the beginning of 1997 to the end of 2003. Because of the

difference between the holidays of these markets, the length of the three series are slightly different. Note that the period we chose covered the Asian financial crisis, this may be a challenge to our model. As usual, we first check the ACF and PACF for these three series and find that the log-return series of Hang Seng and Nikkei indices contain apparent autoregressive components. From the plot of ACFs and PACFs of the three indices, we found Hang Seng and Nikkei indices should have their linear components removed before model fitting since our model only focuses on the residuals. After removing the autoregressive components, we fit the residuals to our direct-VaR models. The linear part for Hang Seng index is estimated to be

$$r_t - 0.0381r_{t-1} + 0.0665r_{t-2} - 0.1124r_{t-3} + 0.061r_{t-4} = x_t. \tag{26}$$

Then we fit to x_t a TGARCH direct-VaR model:

$$\mathrm{VaR}_t^2 = \begin{cases} 1.35 \times 10^{-5} + 0.127x_{t-1}^2 + 0.712\,\mathrm{VaR}_{t-1}^2, & x_{t-1} \leqslant 1.126 \times 10^{-2}, \\ 1.42 \times 10^{-5} + 0.229x_{t-1}^2 + 0.771\,\mathrm{VaR}_{t-1}^2, & x_{t-1} > 1.126 \times 10^{-2}. \end{cases} \tag{27}$$

We use the last 500 residuals to check the performance of our model. We do not update our model when new data points feed in and find that there are 21 points outside the 5% VaR. This amounts to an empirical coverage of 4.2%. We also fit to the x_t a mixture GARCH direct-VaR model:

$$\begin{aligned} \mathrm{VaR}_t^2 = {} & z_t(2.03 \times 10^{-4} + 0.141x_{t-1}^2 + 0.671\,\mathrm{VaR}_{t-1}^2) \\ & + (1 - z_t)(1.33 \times 10^{-5} + 0.213x_{t-1}^2 + 0.692\,\mathrm{VaR}_{t-1}^2). \end{aligned} \tag{28}$$

The Bernoulli variable $z_t = 1$ with probability 0.2895. Similar to TGARCH, the out-sample checking shows that there are 29 points outside the 5% VaR. This amounts to an empirical coverage of 5.8%.

The Nikkei index also shows a significant autoregressive component. The linear component is estimated to be

$$r_t + 0.0462 r_{t-1} + 0.687 r_{t-2} = x_t. \tag{29}$$

Fitting to the residuals x_t a TGARCH direct-VaR model gives:

$$\mathrm{VaR}_t^2 = \begin{cases} 3.02 \times 10^{-5} + 1.90 \times 10^{-1} x_{t-1}^2 \\ \quad +8.08 \times 10^{-1}\, \mathrm{VaR}_{t-1}^2, & x_{t-1} \leqslant -5.43 \times 10^{-2}, \\ 6.02 \times 10^{-5} + 2.13 \times 10^{-1} x_{t-1}^2 \\ \quad +6.43 \times 10^{-1}\, \mathrm{VaR}_{t-1}^2, & x_{t-1} > -5.43 \times 10^{-2}, \end{cases} \tag{30}$$

and we find that there are 22 points lying outside the 5% VaR, about 4.4% coverage. If we fit to the residuals x_t a mixture-GARCH direct-VaR model, we got,

$$\begin{aligned} \mathrm{VaR}_t^2 &= z_t(4.127 \times 10^{-5} + 0.168 x_{t-1}^2 + 0.687\, \mathrm{VaR}_{t-1}^2) \\ &\quad + (1 - z_t)(4.067 \times 10^{-5} + 0.315 x_{t-1}^2 + 0.415\, \mathrm{VaR}_{t-1}^2), \end{aligned} \tag{31}$$

where $z_t = 1$ with probability 0.4129 and find that 23 observations lie outside the 5% VaR, about 4.6% empirical coverage.

The ACF of the Nasdaq does not show any apparent pattern. Therefore, we fit to the log-return our TGARCH and mixture GARCH direct-VaR models directly.

The TGARCH direct-VaR model is

$$\mathrm{VaR}_t^2 = \begin{cases} 1.36 \times 10^{-5} + 2.13 \times 10^{-1} r_{t-1}^2 & \\ \quad +8.76 \times 10^{-1}\,\mathrm{VaR}_{t-1}^2, & r_{t-1} \leqslant -1.925 \times 10^{-3}, \\ 1.24 \times 10^{-5} + 1.791 \times 10^{-1} r_{t-1}^2 & \\ \quad +7.32 \times 10^{-1}\,\mathrm{VaR}_{t-1}^2, & r_{t-1} > -1.925 \times 10^{-3}. \end{cases} \tag{32}$$

There are 24 points outside the 5% VaR which amount to a 4.8% coverage. The mixture GARCH model for Nasdaq is

$$\begin{aligned} \mathrm{VaR}_t^2 &= z_t(1.337 \times 10^{-5} + 0.2018 r_{t-1}^2 + 0.5279\,\mathrm{VaR}_{t-1}^2) \\ &\quad + (1 - z_t)(1.309 \times 10^{-5} + 0.1364 r_{t-1}^2 + 0.6183\,\mathrm{VaR}_{t-1}^2), \end{aligned} \tag{33}$$

where $z_t = 1$ with probability 0.4220. There are 27 points of the last 500 out-sample lying outside the 5% VaR, about 5.4% empirical coverage.

For comparison purpose, we also fit some traditional VaR models for all these series. These models are GARCH model with normal innovation, GARCH model with student-t innovation, integrated GARCH (IGARCH) model with normal innovation, IGARCH model with student-t innovation and the popular RiskMetrics method. We summarize the results in Table 2. From the table, in terms of coverage accuracy, the mixture direct VaR models is among the best for two indices (S&P 500 and Nikkei) and second best in the other two indices (HSI and Nasdaq). The threshold direct VaR is the second best in two indices (HSI and Nikkei) but is the best in the case of Nasdaq. The RiskMetrics is better than the mixture direct VaR model in one index but is inferior in three other cases. The IGARCH

model ties with the mixture direct VaR in one index (Nikkei) but performs not as well in the other three. On the other hand, the GARCH-type models perform the worst in all cases except that the GARCH with normal innovation performs the second best in the Nasdaq case. Table 3 shows the ranks of the results in Table 2. In the case of ties, the average of the respective ranks is used. The smaller the sum of ranks the better is the performance. According to the sum of ranks, it seems that the mixture direct VaR model (sum of ranks is 8.5) gives the most stable and best performance in terms of coverage accuracy and it could be a valuable toolkit for financial risk managers. The threshold direct VaR model performs the second best (sum of ranks is 11.5) and can be another good choice for risk management.

Risk managers also like to calculate the VaR at the 1 % level. Table 4 shows the 1 % results obtained by reestimating all the models at the 1 % level. Table 5 shows the ranks of the results at the 1% level. From the table, in terms of coverage accuracy, the mixture direct VaR models are still among the best for two indices (S&P and HSI). The average performance of the mixture direct VaR models according to the sum of the ranks is still superior to the other models. Since the data points lying outside the 1 % VaR are very few, missing one datum point could change the accuracy dramatically. Therefore, in the 1 % VaR case, that the performance of the threshold direct VaR model is not outstanding does not mean this model is not good. Note that ranks of the threshold direct VaR model are stable, this model could still be a good choice for risk management.

7 Conclusion

In this paper, we extended the CAViaR model of Engle and Manganelli [2004] using a threshold or a mixture approach respectively. The quantile regression method is used in the estimation for TGARCH and mixture-GARCH direct-VaR models. A full modelling methodology is suggested in both the threshold and mixture cases. From the results of the simulations and the post-sample prediction results in the real-data studies, we find that these new extensions are potentially useful. One advantage is that we can use the approaches to study the nonlinear phenomenon due to positive and negative shocks in the financial market and the performance of our model appears to be more stable than other methods.

Some econometricians have pointed out that VaR fails to satisfy the sub-additive property and suggested using the expected shortfall (ES) as an alternative tool for risk measure and management. Artzner et al. [1997] defined ES analytically:

$$E_t(y_t|y_t < q_t, \theta). \tag{34}$$

This is just a conditional expectation of the lower tail. Our method can be extended to the estimation of ES. This will be our agenda for the next paper.

8 Acknowledgments

W.K. Li would like to thank the Croucher Foundation for awarding a Senior Research Fellowship (2003-2004).

References

P. Artzner, F. Delbaen, J. Eber, and D. Heath. Thinking Coherently. *Risk*, 10 (11):68– 71, 1997.

T. Bollerslev. Generalized Autoregressive Conditional Heteroscedasticity. *Review of Economics and Statistics*, 72:498– 505, 1986.

W. H. Chan and M. J. Maheu. Conditional Jump Dynamics in Stock Market Returns. *Journal of Business & Economic Statistics*, 20(3):337– 389, 2002.

J. Danielsson and C. de Vries. Tail Index and Quantile Estimation with Very High Frequency Data. *Journal of Empirical finance*, 4:241– 257, 1997.

W. F. M. DeBondt and R. H. Thaler. Does the Stock Market Overact? *Journal of Finance*, 40:793– 805, 1985.

R. Engle and S. Manganelli. CAViaR: Conditional Value at Risk by Regression Quantiles. *Journal of Business & Economic Statistics*, 22(367– 381), 2004.

R. F. Engle. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflations. *Econometrica*, 50:987– 1007, 1982.

R. F. Engle and T. Bollerslev. Modelling the Persistence of Conditional Variances. *Econometric Reviews*, 5:1– 50, 1986.

K. R. French, W. Schwert, and R. F. Stambaugh. Expected Stock Returns and Volatility. *Journal of Financial Economics*, 19:3– 30, 1987.

J. C. Hull. *Fundamentals of Futures and Options Markets*. Prentice Hall, Upper Saddle River, New Jersey, fourth edition, 2002.

S. Jin. *Nonlinear Time Series Modelling with Applications to Finance and Other Fields*. PhD thesis, The University of Hong Kong, 2005.

P. Jorion. *Value at Risk: the New Benchmark for Managing Financial Risk*. McGraw-Hill, New York, second edition, 2001.

R. Koenker and G. Bassett. Regression Quantiles. *Econometrica*, 46:33– 50, 1978.

C. W. Li and W. K. Li. On a Double-Threshold Autoregressive Heteroscedastic Time Series Model. *Journal of Applied Econometrics*, 11:253– 274, 1996.

J. Liu, W. K. Li, and C. W. Li. On a Threshold Autoregression with Conditional Heteroscedastic Variance. *Journal of Statistical Planning and Inferences*, 62: 279– 300, 1997.

S. Manganelli and R. F. Engle. Value at Risk Models in Finance. Technical report, European Central Bank working paper series No. 75, 2001.

K. Nam, C. S. Pyun, and A. C. Arize. Asymmetric Mean-Reversion and Contrarian Profits: ANST-GARCH Approach. *Journal of Empirical Finance*, 9: 563– 588, 2002.

D. B. Nelson. Conditional Heteroskedasticity in Asset Returns: a New Approach. *Econometrica*, 59:347– 370, 1991.

Basel Committee on Banking Supervision. *Amendment to the Capital Accord to Incorporate Market Risk*. Basle: Bank for International Settlements, 1996.

R. Rabemananjara and Zakoïan. Threshold ARCH Models and Asymmetries in Volatility. *Journal of Applied Econometrics*, 8:31– 49, 1993.

G. W. Schwert and R. F. Stambaugh. Expected Stock Returns and Volatility. *Journal of Financial Economics*, 19:3– 29, 1987.

H. Tong and K. S. Lim. Threshold Autoregressive, Limit Cycles and Cyclical Data. *Journal of the Royal Statistical Society: Series B*, 42:245– 292, 1980.

R. S. Tsay. Testing and Modeling Threshold Autoregressive Processes. *Journal of the American Statistical Association*, 84:231– 240, 1989.

H. White. *Estimation, Inference and Specification Analysis*. Cambridge University Press, New York, 1994.

C. M. Wong and M. K. P. So. On Conditional Moments of GARCH Models with Applications to Multiple Period Value at Risk Estimation. *Statistica Sinica*, 13(4):1015– 1044, 2003.

C. S. Wong and W. K. Li. On a Mixture Autoregressive Model. *Journal of the Royal Statistical Society: Series B*, 62:95– 115, 2000.

C. S. Wong and W. K. Li. On a Mixture Autoregressive Conditional Heteroscedastic Model. *Journal of the American Statistical Association*, 96:982– 995, 2001.

G. Wu and Z. Xiao. An Analysis of Risk Measures. *Journal of Risk*, 4(4), 2002.

J. M. Zakoïan. Threshold Heteroskedastic Models. *Journal of Economics Dynamics and Control*, 18:931– 956, 1994.

Table 1: 5% Quantile Regression Results of Two Models

Model	Parameter	True Value	Mean of Estimation(std. dev.)	Loss greater than VaR
	$a_0^{(1)}$	$0.2 \times 1.64^2 = 0.5279$	0.5028(0.0456)	
	$a_1^{(1)}$	$0.25 \times 1.64^2 = 0.6724$	0.6802(0.0426)	
TGARCH	$b_1^{(1)}$	0.3	0.2820(0.0624)	
(1,1;1,1)	$a_0^{(2)}$	$0.1 \times 1.64^2 = 0.2689$	0.2438(0.0550)	
Direct-VaR	$a_1^{(2)}$	$0.15 \times 1.64^2 = 0.4034$	0.4112(0.0457)	
	$b_1^{(2)}$	0.4	0.3916(0.0678)	
	d	1	1.6946(0.8118)	
	γ	0.0	0.2217(0.1714)	72(4.8%)
	a_{10}	$0.1 \times 1.64^2 = 0.2689$	0.3729(0.0514)	
Mixture	a_{11}	$0.1 \times 1.64^2 = 0.2689$	0.3122(0.0579)	
-GARCH	b_{11}	0.4	0.3678(0.0714)	
Direct	a_{20}	$0.5 \times 1.64^2 = 1.3448$	1.5723(0.0417)	
-VaR	a_{21}	$0.2 \times 1.64^2 = 0.5379$	0.6012(0.0479)	
	b_{21}	0.3	0.2785(0.0372)	
	π	0.3	0.2603(0.0819)	76(5.1%)

Table 2: Results for some commonly used models for the 5 % VaR

	GARCH	GARCH(t)	IGARCH	IGARCH(t)
S&P 500	3.2%	2.8%(11.696)	4.0%	3.6%(10.684)
HSI	3.2%	3.4%(6.882)	3.4%	4.0%(6.127)
Nikkei	5.8%	6.0%(11.038)	4.6%	4.6%(8.977)
Nasdaq	5.4%	5.6%(12.853)	5.4%	5.6%(12.500)

	RiskMetrics	Threshold direct-VaR	Mixture direct-VaR model
S&P 500	4.2%	6.2%	5.6%
HSI	5.6%	4.2%	5.8%
Nikkei	5.8%	4.4%	4.6%
Nasdaq	4.0%	4.8%	5.4%

Inside the parentheses are the degrees of freedom of the student-t distributions.

Table 3: Ranks for some commonly used models for the 5 % VaR

	GARCH	GARCH(t)	IGARCH	IGARCH(t)
S&P 500	6	7	3	5
HSI	7	5.5	5.5	4
Nikkei	5.5	7	2	2
Nasdaq	3	5.5	3	5.5
Sum	21.5	25	13.5	16.5

	RiskMetrics	Threshold direct-VaR	Mixture direct-VaR model
S&P 500	2	4	1
HSI	1	2.5	2.5
Nikkei	5.5	4	2
Nasdaq	7	1	3
Sum	15.5	11.5	8.5

Table 4: Results for some commonly used models for the 1 % VaR

	GARCH	GARCH(t)	IGARCH	IGARCH(t)
S&P 500	0.60%	0.40%(11.696)	0.80%	0.80%(10.684)
HSI	0.60%	0.60%(6.882)	0.60%	0.80%(6.127)
Nikkei	1.60%	1.20%(11.038)	0.80%	1.20%(8.977)
Nasdaq	1.60%	1.00%(12.853)	1.00%	1.00%(12.500)

	RiskMetrics	Threshold direct-VaR	Mixture direct-VaR
S&P 500	0.80%	1.20%	1.00%
HSI	1.20%	0.80%	1.00%
Nikkei	1.20%	0.80%	0.80%
Nasdaq	0.80%	0.80%	0.80%

Inside the parentheses are the degrees of freedom of the student-t distributions.

Table 5: Ranks for some commonly used models for the 1 % VaR

	GARCH	GARCH(t)	IGARCH	IGARCH(t)
S&P 500	6	7	3.5	3.5
HSI	6	6	6	3
Nikkei	7	3.5	3.5	3.5
Nasdaq	7	2	2	2
Sum	26	18.5	15	12

	RiskMetrics	Threshold direct-VaR	Mixture direct-VaR
S&P 500	3.5	3.5	1
HSI	3	3	1
Nikkei	3.5	3.5	3.5
Nasdaq	5	5	5
Sum	15	15	10.5

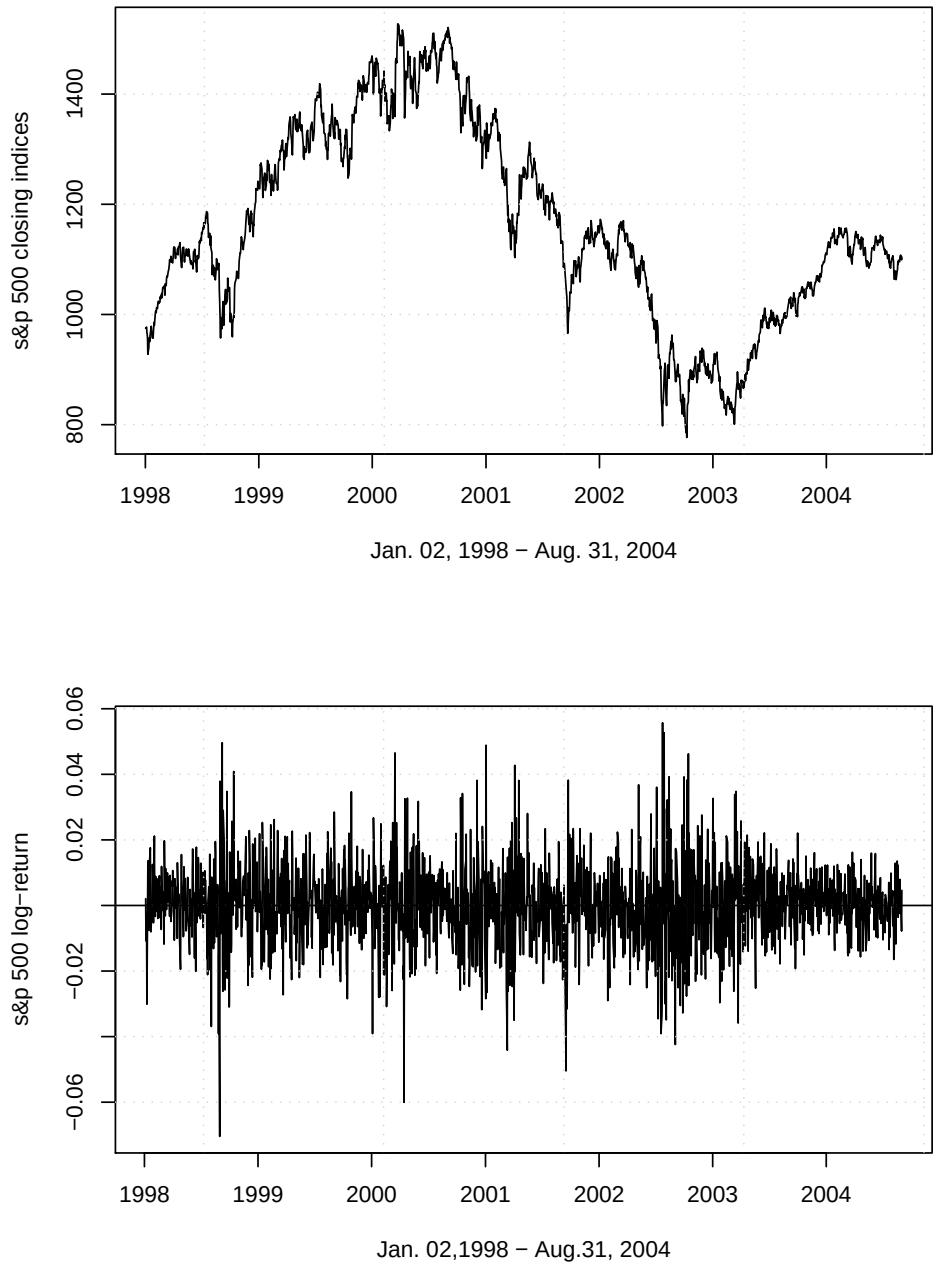

Figure 1: Plot of S&P 500 (Jan. 2, 1998 – Aug. 31, 2004). The upper panel is the plot of the real series and the lower panel is the plot of the log-return series.

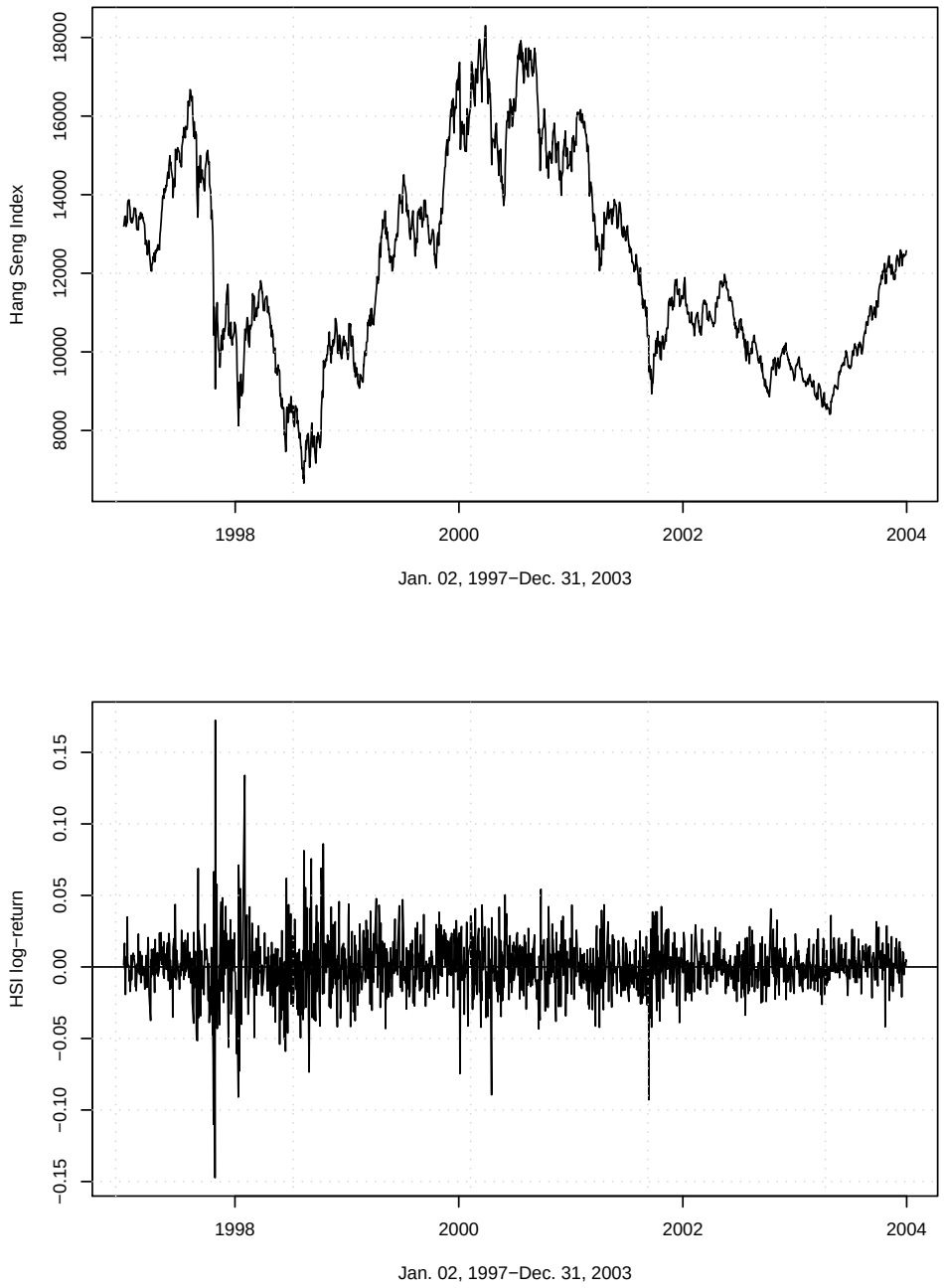

Figure 2: Plot of Hang Seng Index (Jan. 2, 1997 — Dec. 31, 2003). The upper panel is the plot of the real series and the lower panel is the plot of the log-return series.