Forecasting Value-at-Risk Using the Markov-Switching ARCH Model

Yin-Feng Gau*

Wei-Ting Tang

Department of International Business

National Chi Nan University

Abstract

This paper analyzes the application of the Markov-switching ARCH model (Hamilton and Susmel, 1994) in improving value-at-risk (VaR) forecast. By considering a mixture of normal distributions with varying variances over different time and regimes, we find that the "spurious high persistence" found in the GARCH model is adjusted. Under relative performance and hypothesis-testing evaluations, the VaR forecasts derived from the Markov-switching ARCH model are preferred to alternative parametric and nonparametric VaR models that only consider time-varying volatility.

JEL classification: C22, C52, G28.

Keywords: Value-at-Risk, Switching-regime ARCH models.

* Corresponding author. Address: Department of International Business, National Chi Nan University, 1 University Rd., Puli, Nontou 545, Taiwan. Tel: +886-49-291-0960 ext. 4921, Fax: +886-49-291-2595, E-mail: yfgau@ncnu.edu.tw.

1. Introduction

With the increasing fluctuations in assets prices and severe financial turmoil occurred recently, the issue of risk management has received considerable attentions recently. Since its adoption by the Basel Committee (Basel Committee on Banking Supervision of Bank for International Settlements, 1996), value-at-risk (VaR) has become one of the most widely used tools for measuring the market risk by major trading institutions. VaR is used to quantify the exposure of a portfolio to future market fluctuations.

The purpose of this paper is twofold. First, we consider approaches that allow for the leptokurtosis in the distribution of the portfolio return. Since assuming normality in calculating VaR will result in a systematic under-estimation of the riskiness of the portfolio, especially when returns are heavily fat-tailed. To capture the leptokurtosis many researchers use the GARCH model of Bollerslev (1986) to generate volatility forecast (Duffie and Pan, 1997).[1] However, GARCH forecasts are too high in volatile periods. Hamilton and Susmel (1994) argue that the problem of spurious persistence can be solved after considering regime switches in the volatility. Using the Markov-switching ARCH (SWARCH) model proposed by Hamilton and Susmel (1994), we forecast VaR allowing for regime switches in time-varying conditional variance of returns.[2] Second, we evaluate VaR forecasts systematically through relative performance comparison and hypothesis tests on forecast accuracy. While the concept of VaR is simple and attractive, there is no unique approach with VaR implementations adopt. Because a wide variety of alternative models are used in VaR

[1] It can be seen that a mixture of normal distributions with different variances will lead to an overall series that is leptokurtic (Duffie and Pan, 1998).

[2] Cai (1994) also proposed a Markov-ARCH model to incorporate the features of both Hamilton's (1989) switching-regime model and Engle's (1982) ARCH model. Since both models of Cai (1994) and Hamilton and Susmel (1994) aim to integrate Markov Switching model and ARCH model, and the two Markov-switching ARCH models are related in paramaterization (see Cai (1994)), we only estimate the model of Hamilton and Susmel (1994) in this paper without loss of generality.

implementations, it is essential to use systematic evaluation criteria in selecting a preferred VaR model.

This paper undertakes four case studies in model evaluation, including the S&P 500 index, Nikkei 225 index, FTSE 100 index and CAC 40 index, at the 95% and 99% levels significance. The empirical results show that the SWARCH model can solve the problem of "spurious high perisitence" found in the GARCH model and yield a better forecast of VaR.[3] The evaluation results indicate that SWARCH-based VaR forecasts are generally more accurate than those generated by models that only consider time-variation in the conditional volatility, including the EWMA (exponential weighted smoothing average), threshold GARCH (TGARCH) methods and the historical simulation adjusted for time-varying variance.

This paper is organized as follows. Section 2 introduces the evaluation framework for VaR forecasts. Section 3 describes the different models used to derive VaR forecasts. Section 4 compares the results of the empirical investigation of competing models on S&P 500, FTSE100, Nikkei225 and CAC40 indices returns. Section 5 concludes.

2. Evaluation of VaR Forecasts

2.1 Definition of VaR

The concept of VaR is to summarize the worst loss over a target horizon with a given level of confidence. VaR is defined as the maximum loss on a portfolio that can be expected with a certain level of confidence (1-α) over a certain interval of time (*T)*, and can be expressed as:

---

[3] Besides the SWARCH model, there are alternative methods to incorporate both structural change and time-varying stochastic volatility to solve the problem of the excessive GARCH forecasts in volatile periods. For example, Gray (1996) and Ang and Bekaert (2002) extend the specification of SWARCH to the Markov switching GARCH model. However, for the regime-switching GARCH specification one is unable to compute the multi-period ahead volatility forecasts.

$$\Pr\left(r_{t,t+T} < -VaR_{t,t+T}\right) = \alpha \tag{1}$$

where $r_{t,t+T}$ represents the portfolio returns over the T periods in the future, that is, $r_{t,t+T} = \frac{P_{t+T} - P_t}{P_t}$, where P_t is the value of portfolio at time t.

2.2 Evaluation Methods of VaR Forecasts

VaR models are only useful when they predict risk reasonably well. To compare various VaR forecasts, we must check systematically the validity of the evaluation of VaR models through the comparison of predicted and actual loss levels. When a VaR model is perfectly calibrated, the number of realized observations falling outside VaR prediction should be in line with the confidence level. With too many exceptions that exceed the estimated VaR, it means that the model underestimates the risk. This is a major problem because too little capital maybe allocated to risk-taking units. With too few exceptions is also a problem because it leads to excess or inefficient allocation of capital.

Recently, there is a rapidly growing literature on the evaluation of VaR models. One type of methods judges the better VaR forecasts based on the relative performance derived from some loss functions. The other offers a testing framework based on certain theoretical properties of the VaR measures. A key issue about evaluation based on the hypothesis-testing framework is the power of test. If the hypothesis tests exhibit low power, the probability of classifying an inaccurate VaR model as “acceptably accurate” will be high.

2.2.1 Relative Performance

The central concept of these methods is to compare among VaR models and select the most accurate one. Hendricks (1996) proposed several criteria to examine different VaR measures. He emphasized that these considered performance criteria do not have straightforward standard error that it is not possible to discriminate between methods using formal statistical hypothesis. Nevertheless, these criteria provide a relatively complete picture of the performance of selected VaR estimates.

Lopez (1998) proposed a measure of relative performance that can be used to monitor the performance of VaR estimates. The general form of a loss function is

$$C_{i,T|t} = \begin{cases} f(r_{t,t+T}, VaR^{(i)}_{t,t+T}) & if\ r_{t,t+T} < -VaR^{(i)}_{t,t+T} \\ g(r_{t,t+T}, VaR^{(i)}_{t,t+T}) & if\ r_{t,t+T} \geq -VaR^{(i)}_{t,t+T} \end{cases} \quad (2)$$

where $C_{i,T|t}$ represents the numerical scores generated for individual VaR model *i*, and $r_{t,t+T}$ represents the portfolio returns over the T days in the future. The score for the complete regulatory sample of size h is $C_i = \sum_{t=s+1}^{s+h} C_{i,T|t}$. Once a loss function is defined and C_i is calculated, a benchmark can be constructed and used to evaluate the performance of a set of VaR forecasts. In this paper, we apply the following five criteria to evaluate the relative performance of various VaR forecasts.

(1) Mean Relative Bias (MRB)

MRB examines whether different VaR models produce similar forecasts. We first calculate VaR under each VaR models on each sample date, and then compute the average VaR over the forecast sample. Given *h* forecasting periods and *N* VaR models, the MRB of model *i* is computed as:

$$MRB_i = \frac{1}{h}\sum_{t=s+1}^{s+h} \frac{VaR^{(i)}_{t,t+T} - \overline{VaR_{t,t+T}}}{\overline{VaR_{t,t+T}}}, \quad \text{where } \overline{VaR_{t,t+T}} = \frac{1}{N}\sum_{i=1}^{N} VaR^{(i)}_{t,t+T} \quad (3)$$

(2) Root Mean Square Relative Bias (RMSRB)

RMSRB examines the degree to which certain VaR measure varies from the average risk measure for a given date. It captures two effects: the extent to which the average risk estimate provided by a given model systematically differs from the average risk measure, and the variability of each model's risk estimate. The RMSRB is computed as:

$$RMSRB_i = \sqrt{\frac{1}{h}\sum_{t=s+1}^{s+h}\left(\frac{VaR_{t,t+T}^{(i)} - \overline{VaR_{t,t+T}}}{\overline{VaR_{t,t+T}}}\right)^2}, \text{ where } \overline{VaR_{t,t+T}} = \frac{1}{N}\sum_{i=1}^{N} VaR_{t,t+T}^{(i)} \quad (4)$$

(3) Correlation between Risk Measure and Absolute Value of Outcome

A simple efficiency test is to measure the correlation between calculated VaR and the absolute value of realized return. It assesses how well the risk measures adjust over time to underlying changes in risk. This correlation statistic has two advantages. First, it is not affected by the scale of the portfolio. Second, the correlations are relatively easy to interpret.

(4) Binary Loss Function

The loss function implied by the binomial method is

$$C_{i,T|t} = \begin{cases} 1 & if\ r_{t,t+T} < -VaR_{t,t+T}^{(i)} \\ 0 & if\ r_{t,t+T} \geq -VaR_{t,t+T}^{(i)} \end{cases} \quad (5)$$

If a loss exceeding the VaR is observed, this is termed an "exception." Here, we are simply concerned with the number of exceptions rather than the magnitude of these exceptions. If a VaR model is truly providing the level of coverage defined by its confidence level, the score for the complete regulatory sample $C_i = \sum_{t=s+1}^{s+h} C_{i,T|t}$ will equal $0.05 \times h$ and $0.01 \times h$ for the 95th percentile VaR and the 99th percentile VaR, respectively.

(5) Quadratic Loss Function

Quadratic loss function takes account of the magnitude of the exceptions. Comparing with a binary loss function, an additional quadratic term is imposed when an exception occurs. Lopez (1998) found that the use of the additional information embodied in the size of the exception provides a more powerful measure of model accuracy than the binary loss function. The loss function is defined by:

$$C_{i,T|t} = \begin{cases} 1 + \left(\left| r_{t,t+T} \right| - VaR_{t,t+T}^{(i)} \right)^2 & if\ r_{t,t+T} < -VaR_{t,t+T}^{(i)} \\ 0 & if\ r_{t,t+T} \geq -VaR_{t,t+T}^{(i)} \end{cases} \quad (6)$$

In the same manner, we compute the score of the quadratic loss function as $C_i = \sum_{t=s+1}^{s+h} C_{i,T|t}$. When the score of the binary loss function is similar under different models, the quadratic loss function goes in depth to examine the magnitude of these exceptions.

2.2.2 Hypothesis-Testing Framework

Evaluation methods based on a hypothesis-testing framework allow us to test the null hypothesis that VaR forecasts are "acceptably accurate." The null hypothesis is that VaR forecasts in question exhibit a specified property or characteristic of accurate VaR forecasts (Lopez, 1998). If the null hypothesis is rejected, the VaR forecasts do not exhibit the specified property, and the underlying VaR model can be said to be "inaccurate." If the null hypothesis cannot be rejected, the model is said to be "acceptably accurate."

Kupiec (1995) is the first one to develop the performance-based verification techniques to test the accuracy of VaR forecasts. He constructed VaR verification tests from the series of Bernoulli trial outcomes generated by a daily performance comparison. That is, treat the loss on trading activities less than the VaR estimated as a success, and beyond the VaR as a failure. According to this assumption, he derived the TUFF (Time Until First Failure) and PF (Proportion of Failures) tests.

In a performance-based verification scheme, the initial monitoring statistic of interest is the number of observations until a failure is observed. Kupiec (1995) defined $\tilde{T}$ as a random variable that denotes the number of days until the first failure is recorded. If p is the probability of a failure on any given day, the probability of observing the first failure in period V is given by:

$$\Pr(\tilde{T} = V) = p(1-p)^{V-1} \quad (7)$$

where $\tilde{T}$ has a geometric distribution with an expected value of (1/p). For example, when $p = 0.01$, the average time until the first failure is 100. Given a realization for $\tilde{T}$, the likelihood ratio (LR) statistic for testing the null hypothesis $p = p^*$ is given :

$$\mathrm{LR}(V, p^*) = -2\mathrm{Log}\{p^*(1-p^*)^{V-1}\} + 2\mathrm{Log}\{(1/V)(1-1/V)^{V-1}\} \quad (8)$$

Under the null hypothesis, $\mathrm{LR}(V, p^*)$ has a chi-square distribution with 1 degree of freedom. According to Kupiec (1995), when testing $p^* = 0.01$, the TUFF(0.05) critical values for V are 6 and 439. That is, if the first failure occurs before the seventh trading day, it can be concluded that $p > 0.01$. If the first failure occurs after the 438th trading day, it can be concluded that $p < 0.01$. Yet it has been suggested that the TUFF statistics has poor ability to distinguish reliably between alternative underlying values for the tail probability associated with a VaR forecast.

The PF test is used to compare the total number of failures observed to the total accumulated sample size. The PF test is based on the proportion of failures in the sample. When the TUFF test cannot reject the null hypothesis, continued monitoring beyond an observed failure will clearly add information that can be used to verify potential loss estimates. The probability of observing x failures in the sample of size h is:

$$\Pr(h, x) = (1-p)^{h-x} p^x \sim binomial(h, x) \quad (9)$$

where p is the probability of a failure on any one of the independent trails.

The LR statistic is given by:

$$-2 \cdot \log\{(1-p^*)^{h-x}(p^*)^x\} + 2 \cdot \log\{(1-[x/h])^{h-x}(x/h)^x\} \quad (10)$$

Under the null hypothesis, $p = p^*$, the PF test has a chi-square distribution with 1 degree of freedom. In a daily monitoring scheme, the PF test is used to compare the total number of failures observed to the total accumulated sample size. Like the TUFF test, the PF test has poor power in small samples. Kupiec (1995) concluded that sample performance-based VaR verification tests require large samples to produce a reliable accuracy assessment.

The Basel rules for backtesting the internal models approach are derived directly from this failure rate test. The Basel Committee has decided that up to four exceptions are acceptable, which defines a "green" zone for the bank. If the number of exceptions is more, the bank falls into a "yellow" or "red" zone and incurs a progressive penalty.

3. Calculation of VaR Forecasts

To calculate VaR, we need further information regarding the distribution of future return $r_{t,t+T}$. Since VaR is equal to the appropriate quantile of the distribution of future portfolio returns, the task of VaR calculation is to estimate the quantile. By focusing on the quantile or extreme value directly, several approaches have employed the quantile regression or the extreme value analysis to calculate VaR directly, including Engle and Manganelli (1999), and Longin (2000) among others.[4] On the other hand, several approaches estimate the full distribution of portfolio returns and then calculate the corresponding quantile as VaR. Depending on the parameterization, approaches to calculating VaR via the whole distribution can be characterized as parametric and nonparametric methods. The parametric, or namely the variance-covariance or factor approach, involves specifying a parametric distribution and estimating the parameters with historical data. Based on the estimated distribution, often assuming normality, one can calculate the appropriate quantile either

[4] For example, Engle and Manganelli (1999) proposed the CAViaR (Conditional Autoregressive Value-at-Risk) model to study the evolution of the quantile over time. They specified a special type of autoregressive process for the conditional quantile. One disadvantage of this model is that it requires the specification of a dynamic equation for the conditional quantile and its validness is subject to misspecification errors. Instead of undertaking the approach of quantile regression or extreme value analysis, we consider an appropriate model to model the conditional volatility of returns, allowing for jumps or regime changes and time-varying volatility at the same time.

analytically or numerically. On the other hand, the nonparametric or portfolio approach involves constructing or simulating the distribution of portfolio returns that mimic the past performance of the portfolio.

3.1 Parametric Models

Parametric models are the most popular models for calculating VaR, and the normality of returns is usually assumed. Under the assumption of normality of daily portfolio returns, $r_{p,t} \sim N(\mu(p)\ ,\sigma^2(p))$, where $\mu(p)$ and $\sigma^2(p)$ are the mean and variance of $r_{p,t}$ respectively, the value of VaR can be calculated by a multiple of the standard deviation of the portfolio returns. That is,

$$VaR_{t,t+T} = C_\alpha \cdot \sigma(p)_{t,t+T} \tag{11}$$

where C_α is the constant that gives the appropriate one-tailed confidence interval, at the $(1-\alpha)$ confidence level, for the standard normal distribution, while $\sigma(p)_{t,t+T}$ is the standard deviation of portfolio returns over the chosen time horizon, T.

3.1.1 Time-Varying Volatility

When implementing the parametric methods to obtain VaR forecasts, we need to forecast $\sigma(p)_{t,t+T}$ at first. While the assumption of normality simplifies the calculation of VaR, it may lead to an inaccurate VaR. If portfolio returns are leptokurtic,[5] the normal distribution will significantly underestimate the likelihood of extreme returns, and so the estimated VaR of the portfolio will generally be too low.

One cause of leptokurtosis in the unconditional distribution of returns is volatility clustering or time-varying volatility. Duffie and Pan (1997) identified the

---

[5] The daily changes in many variables exhibit significant amounts of positive excess kurtosis (Hull and White, 1998). Duffie and Pan (1997) found that S&P 500 daily returns for 1986 to 1996 have an extremely high sample kurtosis of 111, while the kurtosis of a normal distributed shock is 3. These "fat tails" are particularly worrisome precisely because VaR attempts to capture the behavior of the portfolio return in the left tail. In this situation, a model based on a normal distribution would underestimate the proportion of outliers and the true VaR.

empirical volatility of historical data is changing over time in some persistent manner. As Engle (1982) suggested, if returns are normally distributed with time-varying conditional variance, then the unconditional distribution of returns will have tails that are fatter than those of the normal distribution. To allow for time-varying volatility, the parametric approach is typically modified with a model for the conditional variance of returns[6], such as an exponentially weighted moving average (EWMA) or generalized conditional heteroskedastic (GARCH) model (Bollerslev, 1986). Both models specify the current variance of returns as a function of the lagged variance and lagged squared returns.

The RiskMetrics model (J.P. Morgan/Reuters (1996)) proposes the exponentially weighted moving average (EWMA) model to estimate $\sigma(p)_{t,t+T}$:

$$\sigma_t^2 = (1-\lambda)r_{t-1}^2 + \lambda\sigma_{t-1}^2 \quad (12)$$

where λ is the decay factor which is chosen arbitrary by used and is usually taken the value of 0.94 for daily data.

The development of volatility models for measuring and forecasting volatility dynamics began with the ARCH model proposed by Engle (1982). The ARCH model is useful to estimate the variance of r_t conditional on Ω_{t-1}, the information set available at time t-1. The ARCH (q) model is written as:

$$\sigma_t^2 = w + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \quad (13)$$

where $\varepsilon_t = r_t - E(r_t|\Omega_{t-1})$, $\varepsilon_t \mid \Omega_{t-1} \sim N(0,\sigma_t^2)$, and $E(r_t \mid \Omega_{t-1})$ is the conditional mean of r_t.

6 An alternative approach to obtaining volatility forecasts is the implied volatility approach. The implied volatility is derived from matching trading prices of options and an option pricing formula, for example, Black and Scholes (1976). The implied volatility reflects the market opinion on the volatility of asset returns. However, this approach requires more inputs than the history of returns. Therefore, we only discuss volatility models that only require past return.

Bollerslev (1986) extended the ARCH model to the GARCH model. The GARCH model assumes that the conditional variance depends on the latest innovation but also on previous conditional variance. A GARCH model is the more general form for estimating volatility. The representation of the GARCH (p,q) model is:

$$\sigma_t^2 = w + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{14}$$

where $\varepsilon_t = r_t - E(r_t | \Omega_{t-1})$, $\varepsilon_t | \Omega_{t-1} \sim N(0, \sigma_t^2)$. The conditional variance equation is a function of three terms: the mean c, news about the volatility from the previous period, measured as lagged squared residual from the mean equation ε_{t-i}^2, and past conditional variance σ_{t-i}^2. To ensure the positive variance and stationarity, it requires that $w > 0$, $\alpha_i \geq 0, i = 1,2,\ldots,q, \beta_i \geq 0, i = 1,2,\ldots,p$, and $\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i < 1$. For the model specification, $E(r_t | \Omega_{t-1})$ is the conditional mean and could be modeled as an AR or MA process when returns are autocorrelated.

Alexander and Leigh (1997) examined the performance of three volatility models: the equally weighted moving average of squared returns, the exponentially weighted moving average, and GARCH models. They concluded that GARCH models give more conservative risk capital estimates, which can more accurately reflect a 1% value at risk measurement.

However, for equities, it is often observed that downward movements in the market are followed by higher volatilities than upward movements of the same magnitude. To account for this phenomenon, the TGARCH (threshold GARCH) models allows for asymmetric impacts of shocks on current volatility. The specification of the TGARCH(1,1) model, suggested by Glosten, Jagannathan, and Runkle (1993), is written as:

$$\sigma_t^2 = w + \alpha\varepsilon_{t-1}^2 + \beta\,\sigma_{t-1}^2 + \delta \cdot d_{t-1} \cdot \varepsilon_{t-1}^2 \tag{15}$$

where $\varepsilon_t = r_t - E(r_t|\Omega_{t-1})$, $\varepsilon_t \,|\, \Omega_{t-1} \sim N(0,\sigma_t^2)$, $d_{t-1}=1$ if $\varepsilon_{t-1} < 0$, and $d_{t-1}=0$ otherwise. In this model, good news (when $\varepsilon_{t-1} \geq 0$) and bad news (when $\varepsilon_{t-1} < 0$) have different effects on current conditional variance. By the definition, the impact of good news is α, while the impact of bad news is ($\alpha + \delta$). If $\delta > 0$, we say that a leverage effect exists in that bad news increases volatility. The relation $\delta = 0$ implies that the news impact on the current conditional variance is symmetric.

On the other hand, as noted in Duffie and Pan (1997), one possible source of fat tails is jumps, or significant unexpected discontinuous changes in prices. The jump diffusion model has been treated as a recipe for fat-tailed distributions. The major implication of the jump diffusion model for extreme loss shows up much farther out in the tail. To consider both the time-varying volatility and possibility of jumps in the volatility process, we estimate the class of Markov-switching models for the time-varying volatility, namely Markov-switching ARCH (SWARCH) model, proposed by Hamilton and Susmel (1994), to allow for both time-variation and regime switches in the conditional volatility.

3.1.2 Time-Varying Volatility and Regime Switches

GARCH forecasts are usually too high, especially in periods of high volatility. This is due to the high degree of persistence implied from the GARCH model. The problem of "spuriously high persistence" results in the weak forecasting performance, since the impacts of shocks usually do not last for such a long period. As pointed by Hamilton and Susmel (1994), the spuriously high persistence might be related to structural changes in the variance process. The volatility forecast will be less persistent if we model changes in parameters through a Markov-switching process, as shown in Hamilton and Susmel (1994) and Cai (1994) among others.

The SWARCH models proposed by Hamilton and Susmel (1994) allow the volatility dynamics to change under different states or regimes. That is, parameters in the ARCH(q) process are allowed to be changed in different states. State variable s_t indicates the state that the process is in at time t and it is assumed to follow a Markov chain. That means the probability of state $s_t = j$ will be affected by only the realized state in the last period:

$$\Pr(s_t = j \mid \Omega_{t-1}) = \Pr(s_t = j \mid s_{t-1} = i) = p_{ij}$$

We denote $\sigma_t^2 \sim \text{SWARCH}(K,q)$ if and only if σ_t^2 follows a K states, q-th order Markov-switching ARCH process. The model can be written as

$$\sigma_t^2 = g_{s_t} \times \tilde{\sigma}_t^2 \tag{16}$$

$$\tilde{\sigma}_t^2 = w + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \tag{17}$$

where $\varepsilon_t = r_t - E(r_t \mid \Omega_{t-1})$, $\varepsilon_t \mid \Omega_{t-1} \sim N(0, \sigma_t{}^2)$ as before. That is, $\tilde{\sigma}_t^2 \sim \text{ARCH(q)}$. g_{s_t} is the multiplicative factor that depends on the state s_t. Under this model, $\tilde{\sigma}_t^2$ is multiplied by the constant g_1 when the process is in the state 1 or $s_t = 1$, multiplied by g_2 when $s_t = 2$, and so on.

An extension of the SWARCH model is the SWARCH-L model that captures the leverage effect as the specification of a threshold ARCH model. The process of conditional volatility becomes

$$\sigma_t^2 = g_{s_t} \times \tilde{\sigma}_t^2,$$

$$\tilde{\sigma}_t^2 = w + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \delta \cdot d_{t-1} \cdot \varepsilon_{t-1}^2 \tag{18}$$

where $\varepsilon_t = r_t - E(r_t \mid \Omega_{t-1})$, $\varepsilon_t \mid \Omega_{t-1} \sim N(0, \sigma_t{}^2)$, $d_{t-1} = 1$ if $\varepsilon_{t-1} < 0$, and $d_{t-1} = 0$ otherwise. For parameters estimation and forecasts calculation, please refer to

Hamilton and Susmel (1994) for more details.

3.2 Nonparametric Models

Nonparametric models are independent from the parameterized distribution of assets returns or market factors returns. One of which is the historical simulation method.[7] The method was proposed initially by Efron (1979) as a nonparametric randomization technique that constructs the empirical distribution by drawing from the observed distribution of the data. It simply requires relatively few assumptions about the statistical distributions of the underlying market factors because it assumes that market prices innovations in the future are drawn from the same empirical distribution as those market price innovations generated historically.

Instead of estimating parameters, such as the standard deviation, the method of historical simulation simply uses the actual percentiles of the observation period as VaR measures. This method involves creating a database consisting of the daily movements in all market variables over a period of time. If we assume that the returns in the next day are simply associated with the period of historical observations, we could directly rank the observed historical returns, and apply these ranked historical returns to construct the distribution of return in the next period.[8]

The method of historical simulation requires no parameter in estimating the empirical distribution. However, if the future distribution of market factors differs

---

7 Stress testing is another kind of nonparametric models. The goal of stress testing is to identify unusual scenarios that would not occur under standard VaR models. In some sense, stress testing can be viewed as an extension of the historical simulation method at increasingly higher confidence level. (Jorion, 2000) On the other hand, the Monte Carlo simulation method is an alternative approach of parametric models. The method is used to simulate a variety of different scenarios for the portfolio value on the target date by generating random draws for the risk factors from a predetermined distribution. In a Monte Carlo simulation, one chooses a statistical distribution that is believed to adequately approximate the possible changes in the market factors. Then, a pseudo-random number generator is used to generate thousands (or perhaps tens of thousands) of hypothetical changes in the market factors. These hypothetical changes are used to construct thousands of hypothetical portfolio returns on the current portfolio and the distribution of returns. Finally, the VaR is computed from this distribution.

8 For example, suppose that 1,000 days of data are used and the 1 percentile of the distribution is required. VaR would be estimated as the tenth worst change in the portfolio value.

substantially from the historical distribution, computed results can be misleading. Hull and White (1998) modified the method of historical simulation using an adjustment on the variance. Instead of using the actual historical percentage changes in market variables to calculate VaR, they used historical changes that have been adjusted to reflect the ratio of the current daily volatility to the daily volatility at the time of the observation.

Let r_t be the historical percentage change in the price on day t, a period covered by the historical sample N (that is, $t < N$); σ_t^2 be the historical estimate of the variance of return for day t. Then the most recent estimate of the daily variance is σ_N^2, the variance estimate made at the end of day N-1. Assuming the process of r_t / σ_t^2 is stationary, then the adjusted r_t, r_t^*, is given by: $r_t^* = \sigma_N \frac{r_t}{\sigma_t}$, where σ_t can be the estimated volatility from the historical data. In this paper, we use the TGARCH (threshold GARCH) model to estimate the volatility for adjusting the historical observations.

4. Empirical Results

The data studied in this paper are returns to major stock indices, including S&P 500, FTSE 100, NIKKEI 225, and CAC 40 indices. The daily data are collected from the Datastream and cover the period from January 1990 through December 2002.

We calculate daily log returns by taking the difference of log prices for each index. Table 1 reports the descriptive statistics of these stock index returns. It shows that values of sample mean are close to 0. Values of sample kurtosis lie between 5.945 and 8.764. The kurtosis is higher than 3, the kurtosis of a normal distribution, which shows that distributions of index returns exhibit fat tails. By the Jarque-Bera statistic, the null hypothesis of normal distributions is also rejected for all four returns. We also detect weak first-order autocorrelation in returns to four indices. Values of the Ljung-Box Q statistic suggest the existence of significant serial correlation in returns and squared returns from the four indices. By the phenomenon of autocorrelated squared returns, we see that the data exhibit the characteristic of volatility clustering.

The number of observations for each index return in this study is 3391. We use the last 500 observations for out-of-sample forecasting. The estimation procedure that we apply is as follows. For each model, 2891 observations of daily data are used in estimation, and used to form a VaR forecast for day 2892. After this, data from day 2 until 2892 is used in estimation to obtain a VaR forecast for day 2893. For each of the VaR models competing in this paper, 500 out-of-sample forecasts are generated recursively by moving the estimation-window forward through time.

Assuming that the conditional distribution of returns is normal, we can obtain VaR via the formula: $VaR_{t,t+T} = C_{\alpha} \cdot \sigma_{r,r+T}$, where C_{α} is a multiplicative factor that depends on the confidence level ($1-\alpha$) of a normal distribution, and $\sigma_{t,t+T}$ is the standard deviation of returns over T periods. For the simplest case, we set T=1. Since $VaR_{t,t+1} = C_{\alpha} \cdot \sigma_{t,t+1} = C_{\alpha} \cdot \sigma_{t+1}$, we calculate VaR with $\hat{V}aR_{t,t+1} = C_{\alpha} \cdot \hat{\sigma}_{t+1|t}$, where $\hat{\sigma}_{t+1|t} = \sqrt{\hat{\sigma}^2_{t+1|t}}$ and $\hat{\sigma}^2_{t+1|t}$ is the forecast of Var (r_{t+1}) conditional on information available at date t.

In the EWMA model, $\hat{\sigma}^2_{t+1|t} = (1-\lambda)r_t^2 + \lambda\sigma_t^2$, we set λ to 0.94 as J.P. Morgan suggests. After forecasting the variance, daily VaR is computed as $1.645 \times \hat{\sigma}_{t+1|t}$ for the 95%, and $1.96 \times \hat{\sigma}_{t+1|t}$ for the 99% VaR. However, the Ljung-Box Q statistics for the squared standardized returns show that the squared standardized returns are still autocorrelated when we use the EWMA to estimate the time-varying volatility. This indicates that the EWMA does not capture the time-varying volatility well enough.

On the contrary, we find GARCH (1,1) is sufficient to capture the volatility clustering, and the leverage effect is significant. To capture the leverage effect in index return volatility, we use the TGARCH (1,1) model to forecast the volatility. Table 2 reports the estimation results of the TGARCH(1,1) models for observations 1 to 2891. It indicates that, in the TGARCH (1,1) model, the persistence that can be measured by the sum of ARCH and GARCH parameters, i.e., $(\alpha_1 + \beta_1 + \frac{1}{2}\delta)$, is close to unity for each series. This may indicate spurious high volatility persistence. Furthermore, we use the statistic of CUSUM of squares to roughly examine if there might be structural changes in the variance process. If the hypothesis of no structural change fails to accept, it implies that structural changes may exist and that regime-switching models are appropriate for estimating the volatility process. In this paper, we use the CUSUM of squares test as a simple diagnostic for the stability of the variance process. As with the CUSUM of squares test, movement outside the critical lines is suggestive of parameter or variance instability. According to Figure 1, the results of CUSUM-squares statistic suggest that, for each index return, there exists variance instability.[9] It implies that the "spuriously high persistence" might be related to structural changes. Therefore, we further set a two-regime SWARCH model to estimate and forecast the volatility.[10]

Table 3 reports the estimation results of the SWARCH-L model using

---

[9] A formal statistical test for the null hypothesis of no-regime switching has been proposed in Hansen (1992). In this paper, we are focused on the problem of the excessive GARCH forecasts in volatile periods and thus consider the SWARCH model to allow for regimes with different volatility levels. We use the CUSUM-squares statistic as a simple diagnosis for the possibility of different volatility regimes.

[10] We use two regimes and do no consider models with more regimes, because we want to explore whether the introduction of regimes help solve the "spurious high persistence" problem with the GARCH forecasts and it turns out that two regimes are sufficient for that.

observations 1 to 2891. It shows that the staying probability in each SWARCH-L model is not high (0.4~0.6), especially for the high-volatility regimes. This implies that the duration of high-volatility is not long. That is, the effects of shocks don't always last persistently.

To compare with VaR calculations calculated from parametric volatility models, we implement two alternative nonparametric approaches, including the historical simulation (hereafter, HS) and TGARCH-adjusted historical simulation (adjusted HS).[11] The critical parameter in the HS models is the window width that is used in estimation. We report results for two cases: 500 and 1000 days. The results show that VaR forecasts from the traditional HS approach are fixed for a long period. These forecasts would not change until a great loss occurs or the losses deviate away from the mean of the moving windows. For the case of 500-day window, we use the prior 500 observations to estimate the TGARCH model, and then use the variance forecast for day 2892, the first out-of-forecast observation, to adjust the variance for implementing the adjusted HS. Similarly, we move forward through time, generating out-of-sample adjusted-HS. Table 4 reports the statistical summary of VaR forecasts from competing VaR models In average, except for the FTSE 100, VaR forecasts from the SWARCH-L model are higher than any other measures.

Relative performances of models compared are given in Table 5. For the 95% VaR, SWARCH-L model produces the highest MRB, meaning its VaR forecasts are much higher than the average over all models compared. MRB of the other models are between -0.16 and 0.16, showing that the differences across these models are relatively small. For the 99% VaR, MRB are higher in each model, between -0.20 and 0.29, which shows that the forecasts from these VaR models are not similar.

[11] Following Hull and White's (1998) procedure we implement the HS adjusted by the TGARCH volatility.

While MRB measure the average deviation of a model from the average VaR across models, RMSRB measures the dispersion of a specific VaR model deviating from the average VaR across models. According to Table 4, the results of RMSRB indicate that the SWARCH model produces higher VaR estimations at a given date for both the 95% and 99% VaRs.

The results of correlation between the VaR forecast and the realized return suggest the superior performance of the SWARCH-L model. The SWARCH-L model exhibits the highest correlation among all models for the four index returns examined. Besides, VaR calculated from the TGARCH and TGARCH-adjusted HS models performs relatively better than that from the EWMA and historical simulation models. Unsurprisingly, the HS models have the lowest correlation. This shows that the historical-simulated VaR is unable to well track changes in risk over time.

The benchmark of score based on the binary loss function is 25 for the 95% VaR and 5 for the 99% VaR. For the 95% VaR, We found the score of the SWARCH-L model is lower than 25, and the score of the adjusted HS model is most close to 25. The scores of the EWMA, TGARCH, and HS models are much higher, indicating there are much more exceptions exceeding VaR. For the 99% VaR, the results are similar, but the SWARCH-L and HS (1000 days) models performs the best for the Nikkei 225 returns.

The score based on the quadratic loss function measures the magnitude of exceptions. For the 95% VaR, the score of the SWARCH-L model is the lowest. We believe that it is resulted by its fewer exceptions. We also found that the HS achieves the highest score although it exhibits the same number of exceptions with the EWMA model (for NIKKEI 225) or even fewer exceptions than the TGARCH model (for FTSE 100). As for the 99% VaR, the score of adjusted HS model is relatively lower than the other models.

For the HS models, we use observation periods of 500 and 1,000 days moving window. The performance under different observation periods is not greatly dissimilar. Adjusted-HS model consistently performs better than the historical model across each criterion. The HS model tends to produce higher scores of loss function, implying it underestimates risk. Besides, its estimations have the lowest correlation between actual outcomes among all models. That is, it has poor ability to adjust risk measures over time.

According to the number of actual loss exceeding VaR, we calculate each model's LR statistics for PF test. For the 95% VaR, only historical simulation models for CAC 40 series reject the hypothesis of $\Pr\left(r_{t,t+T} < -VaR_{t,t+T}\right) = 5\%$ at the 99% confidence level. However, for the 99% VaR, the LR statistics of both the EWMA and TGARCH methods extremely exceed the 1% critical value of $\chi^2(1) = 6.63$. Exceptionally, the SWARCH-L and HS models do not perform well for the FTSE 100 series. For the S&P 500, NIKKEI 225, and CAC 40, the PF test cannot reject the hypothesis of $\Pr\left(r_{t,t+T} < -VaR_{t,t+T}\right) = 1\%$ under the SWARCH-L, HS, and adjusted HS models. This is probably because the PF tests generally indicate the coverage probability is correct for most models, especially for the 95% VaR. In summary, the SWRCH-L model tends to produce too few exceptions, although the PF test does not reject its accuracy. The strength of the SWARCH-L model is its efficiency to track the evolution of risk in terms of its highest correlation.

5. Conclusion

This paper evaluates the forecasting performance of the SWARCH model based on a systematic evaluation for the corresponding VaR forecasts. VaR has been widely used to quantify and control the market risk, and the better forecast of volatility help improving the VaR forecasts. The estimation results show that the high degree of persistence estimated from the widely used GARCH models can be adjusted by allowing regime switches in the time-varying volatility.

By evaluating out-of-sample VaR forecasts via relative performances based on certain loss functions and the hypothesis testing based on the LR statistic, we conclude that the SWARCH-L model outperforms alternative competing models, including the RiskMetric or EWMA model, TGARCH model, HS model, and adjusted HS model.

In this paper we are focused on the problem of the excessive GARCH forecasts in volatile periods and thus consider the SWARCH model to allow for regimes with different volatility levels. It is left to the future research to examine if a SWARCH model with more regimes, or a Markov switching GARCH model can explain the dynamics of time-varying volatility better. Besides, we only examine the performance of daily VaR forecasts. Furthermore, Certain institutions, however, care their trading risk under longer holding periods. It is also commendable to evaluate VaR forecasts of each model under different horizons.

References

Alexander, C. O. and C. T. Leigh (1997), "On The Covariance Matrices Used in Value at Risk Models," *Journal of Derivatives*, 4, 50-62.

Ang, A. and G. Bekaert (2002), "Regime Switches in Interest Rates," *Journal of Business and Economic Statistics*, 20, 163-182.

Cai, J. (1994), "A Markov Model of Unconditional Variance in ARCH," *Journal of Business and Economic Statistics*, 12, 309-316.

Duffie, D. and J. Pan (1997), "An Overview of Value at Risk," *Journal of Derivatives*, 4, 7-49.

Efron, B. (1979), "Bootstrapping Methods: Another Look at the Jackknife," *Annals of Statistics*, 7, 1-26.

Engle, R. F. and S. Manganelli (1999), "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," UCSD Discussion Paper.

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

Gray, S. F. (1996), "Modeling the Conditional Distribution of Interest Rates as a Regime-switching Process," *Journal of Financial Economics*, 42, 27-62.

Hamilton, J. D. and R. Susmel (1994), "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics*, 64, 307-333.

Hansen, B. E. (1992), "The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP," *Journal of Applied Econometrics*, 7, S61-S82.

Hendricks, D. (1996), "Evaluation of Value-at-risk Models Using Historical Data," *Federal Reserve Bank of New York Economic Policy Review*, 2, 39-70.

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

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

J.P. Morgan/Reuters (1996), *RiskMetrics – Technical Document* (4th edition), New York.

Jorion, P. (2000), *Value at Risk*, 2nd edition, McGraw-Hill.

Kupiec, P. H. (1995), "Techniques for Verifying The Accuracy of Risk Measurement Models," *Journal of Derivatives*, 3, 73-84.

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

Lopez, J. A. (1998), "Methods for Evaluating Value-at-Risk," *Federal Reserve Bank of New York Economic Policy Review*, 119-124.

Penza, P. and Bansal, V. K. (2001), *Measuring Market Risk with Value at Risk*, John Wiley & Sons, Inc.

Table 1

Descriptive Statistics of Index Returns (%)

Statistics	S&P 500	NIKKEI 225	FTSE 100	CAC 40
Mean	0.027	-0.046	0.014	0.012
Median	0.004	0.000	0.000	0.000
Maximum	5.573	12.430	5.440	10.251
Minimum	-7.113	-7.234	-5.885	-13.378
Std. Dev.	1.034	1.507	1.045	1.463
Skewness	-0.116	0.263	-0.127	-0.298
Kurtosis	7.038	6.473	5.945	8.764
Jarque-Bera	2311.832*	1743.389*	1234.501*	4744.620*
$\rho(1)$	0.010	-0.015	0.062*	-0.057*
Q(15)	43.571*	31.692*	46.688*	31.927*
$Q^2(15)$	478.73*	361.27*	740.47*	486.47*

Jarque-Bera is the Jarque-Bera statistic for normality. $\rho(1)$ indicates the first order autocorrelation in returns. Q(15) and $Q^2(15)$ report values of the Ljung-Box Q statistic for up to 15th-order autocorrelation in return and squared returns, respectively. *: Significant at the 1% level of significance.

Table 2 Estimation Results of the TGARCH(1,1) Model

(Observations 1 to 2891)

$r_t = c_0 + \phi r_{t-1} + \varepsilon_t$, where $\varepsilon_t = r_t - \phi r_{t-1}$, $\varepsilon_t \mid \Omega_{t-1} \sim N(0, \sigma_t^{\,2})$

$\sigma_t^2 = w + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \delta \cdot d_{t-1} \cdot \varepsilon_{t-1}^2$, $d_{t-1} = 1$ if $\varepsilon_{t-1} < 0$, and $d_{t-1} = 0$ otherwise.

Parameter	S&P 500	NIKKEI 225	FTSE 100	CAC 40
c_0	0.034*	-0.040*	0.028*	0.018
	(0.014)	(0.022)	(0.016)	(0.022)
ϕ	0.054**	-0.011	0.053**	-0.003
	(0.019)	(0.020)	(0.019)	(0.022)
w	0.011**	0.053**	0.006**	0.064**
	(0.003)	(0.016)	(0.002)	(0.018)
α	0.010	0.017**	0.013*	0.027**
	(0.010)	(0.002)	(0.006)	(0.002)
β	0.926**	0.898**	0.955**	0.890**
	(0.012)	(0.015)	(0.009)	(0.022)
δ	0.106**	0.125**	0.050**	0.092**
	(0.021)	(0.023)	(0.014)	(0.028)
Q(15)	22.777	8.349	16.692	16.763
$Q^2(15)$	11.122	12.633	14.403	15.128

Q(15) and $Q^2(15)$ indicate the Ljung-Box statistics for upto 15-th order autocorrelation in standardized residuals and squared standardized residuals, respectively.

Bollerslev-Wooldridge robust standard errors are reported in parentheses.

**: Significant at the 1% level of significance.

*: Significant at the 5% level of significance.

Table 3　Estimation Results of the SWARCH-L (2,2) Model

(Observations 1 to 2891)

$r_t = c_0 + \phi r_{t-1} + \varepsilon_t$, where $\varepsilon_t = r_t - E(r_t \mid \Omega_{t-1})$, $\varepsilon_t \mid \Omega_{t-1} \sim N(0, \sigma_t^{\ 2})$

$\sigma_t^2 = g_{s_t} \times \tilde{\sigma}_t^2$, $s_t = 1$ for state 1, $s_t = 2$ for state 2; g_1 is set to equal to 1, and the

transitional probability is $\Pr(s_t = j \mid \Omega_{t-1}) = \Pr(s_t = j \mid s_{t-1} = i) = p_{ij}$, $i, j = 1, 2$

$\tilde{\sigma}_t^2 = w + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \delta \cdot d_{t-1} \cdot \varepsilon_{t-1}^2$, $d_{t-1} = 1$ if $\varepsilon_{t-1} < 0$, and $d_{t-1} = 0$ otherwise.

Parameter	S&P 500	NIKKEI 225	FTSE 100	CAC 40
c_0	0.052**	-0.024	0.037*	0.059**
	(0.015)	(0.034)	(0.015)	(0.023)
ϕ	-0.005	-0.036**	0.054**	-0.037*
	(0.020)	(0.019)	(0.020)	(0.021)
w	0.128**	0.333**	0.329**	0.413**
	(0.026)	(0.088)	(0.034)	(0.077)
α_1	0.135**	0.068*	0.149**	0.160**
	(0.043)	(0.031)	(0.039)	(0.044)
α_2	0.239**	0.275**	0.171**	0.133**
	(0.042)	(0.042)	(0.033)	(0.031)
δ	0.236**	0.257**	0.263**	0.148**
	(0.072)	(0.064)	(0.052)	(0.059)
g_2	5.926**	5.606**	3.176**	4.283**
	(0.711)	(0.913)	(0.254)	(0.466)
$\hat{p}_{11}$	0.406**	0.414**	0.614**	0.414**
	(0.081)	(0.108)	(0.104)	(0.113)
$\hat{p}_{22}$	0.420**	0.434**	0.352**	0.443**
	(0.088)	(0.111)	(0.103)	(0.089)

Standard errors are reported in parentheses.

**: Significant at the 1% level of significance.

*: Significant at the 5% level of significance.

Table 4 VaR Calculations

Index Return	S&P 500		NIKKEI 225		FTSE 100		CAC 40	
95% VaR	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
EWMA	2.2986	0.8640	2.8028	0.5433	2.3292	1.0043	3.0303	1.2726
TGARCH	2.2828	0.6778	2.7433	0.5820	2.3407	0.8836	2.8668	1.0617
SWARCH-L	3.1660	1.7361	3.7666	1.6717	2.1873	0.9002	3.4026	1.8343
HS (500 days)	2.1658	0.1570	2.4447	0.1661	2.1129	0.2223	2.6315	0.2347
HS (1000 days)	2.0984	0.0809	2.4674	0.0248	2.1058	0.0494	2.6240	0.0940
Adjusted HS (500 days)	2.4007	0.7263	2.7801	0.5799	2.6730	1.1174	3.1861	1.2217
Adjusted HS (1000 days)	2.4527	0.7140	2.7747	0.6008	2.5764	1.0089	3.0352	0.0425
95% VaR	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
EWMA	2.7388	0.7252	3.3395	0.6473	2.7753	1.1966	3.6106	1.5163
TGARCH	2.7200	0.8075	3.2686	0.6934	2.7890	1.0528	3.4176	1.2650
SWARCH-L	3.7723	2.0685	4.4878	1.9919	2.6062	1.1798	4.0542	2.1856
HS (500 days)	3.1454	0.2095	3.7829	0.2399	3.4358	0.6623	4.1549	0.6629
HS (1000 days)	3.1622	0.0903	3.8677	0.2740	3.3580	0.3227	4.8048	0.1553
Adjusted HS (500 days)	3.2521	1.0229	4.0981	0.9328	3.6012	1.4407	4.4393	1.5864
Adjusted HS (1000 days)	3.8724	1.0302	4.2664	0.9062	3.8024	1.4292	4.8096	1.7147

Table 5 Performances of VaR Forecasts

Panel A. S&P 500

	MRB	RMSRB	Correlation	Binary Loss Function	Quadratic Loss Function	LR(PF Test)
95%VaR						
EWMA	-0.05243	0.14451	0.32005	34	63.52930	3.08057
TGARCH	-0.05987	0.10086	0.40228	30	53.33674	0.99211
SWARCH-L	0.25397	0.48060	0.48388	17	42.96565	3.02146
HS (500 days)	-0.05835	0.19094	0.20683	32	64.66957	1.90271
HS (1000 days)	-0.08262	0.20841	0.12225	38	75.96927	6.18107
Adjusted HS (500 days)	-0.01217	0.09018	0.40246	23	43.26180	0.17286
Adjusted HS (1000 dsays)	0.01146	0.08605	0.39950	20	39.08240	1.12671
99%VaR						
EWMA	-0.16240	0.20325	0.32005	17	35.63731	17.90165
TGARCH	-0.16908	0.18484	0.40228	15	28.76313	13.16176
SWARCH-L	0.11222	0.39247	0.48388	10	29.40874	3.91362
HS (500 days)	0.01387	0.19378	0.13795	8	16.28288	1.53828
HS (1000 days)	0.02336	0.20562	0.09656	9	17.00353	2.61257
Adjusted HS (500 days)	-0.01024	0.10735	0.39160	8	16.76973	1.53828
Adjusted HS (1000 days)	0.19226	0.22097	0.37877	3	6.18547	0.94312

Panel B. FTSE 100

	MRB	RMSRB	Correlation	Binary Loss Function	Quadratic Loss Function	LR(PF Test)
95%VaR						
EWMA	-0.00516	0.12781	0.10486	33	61.69232	2.45919
TGARCH	-0.03233	0.09861	0.15279	31	56.20068	1.41302
SWARCH-L	0.29622	0.48677	0.30609	14	20.21085	6.01788
Historical (500 days)	-0.11300	0.18373	-0.02644	33	90.52959	2.45919
Historical (1000 days)	-0.10488	0.17154	-0.03438	34	85.72846	3.08057
Adjusted HS (500 days)	-0.01927	0.09841	0.14356	30	55.84852	0.99211
Adjusted HS (1000 days)	-0.02159	0.10058	0.15046	30	53.99998	0.99211
99%VaR						
EWMA	-0.13646	0.17165	0.10486	15	29.25933	13.16176
TGARCH	-0.15983	0.17644	0.15279	15	25.13434	13.16176
SWARCH-L	0.12976	0.37219	0.30609	5	6.21305	0.00000
HS (500 days)	-0.00096	0.15097	0.00462	8	23.72890	1.53828
HS (1000 days)	0.01961	0.15103	0.02438	5	16.93498	0.00000
Adjusted HS (500 days)	0.05134	0.11558	0.15039	3	5.33538	0.94312
Adjusted HS (1000 days)	0.09654	0.13694	0.15415	3	4.41733	0.94312

Table 5 (Continued)

Performances of VaR Forecasts

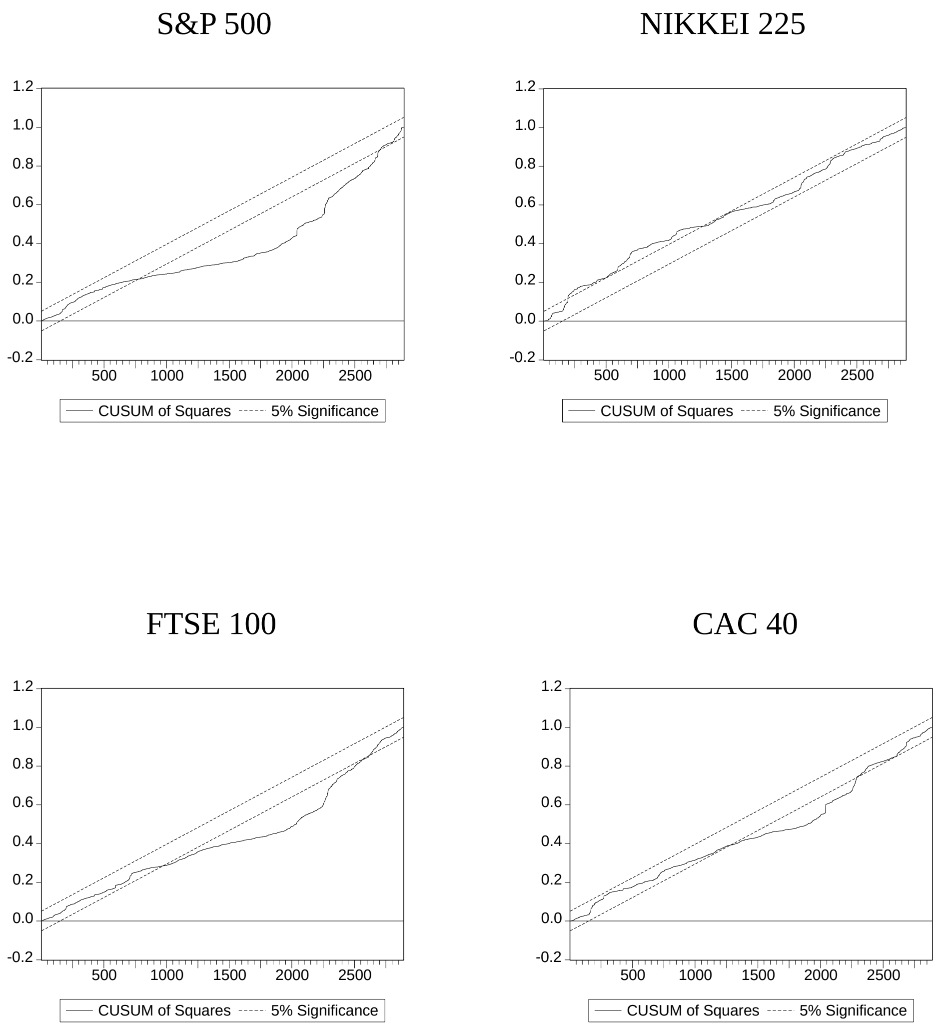
Panel C. NIKKEI 225

	MRB	RMSRB	Correlation	Binary Loss	Quadratic Loss	LR(PF Test)
95%VaR						
EWMA	-0.03688	0.14662	0.41677	40	81.06682	8.07904
TGARCH	-0.01768	0.09052	0.43412	38	78.26163	6.18107
SWARCH-L	-0.07148	0.24108	0.60511	21	46.19797	0.71075
HS (500 days)	-0.03565	0.21197	0.15457	35	131.13609	3.76508
HS (1000 days)	-0.02495	0.25211	0.07962	34	131.01747	3.08057
Adjusted HS (500 dsays)	0.10939	0.17969	0.43507	24	49.70562	0.04265
Adjusted HS (1000 days)	0.07725	0.13432	0.42963	28	57.73407	0.36539
99%VaR						
EWMA	-0.16239	0.20479	0.41677	24	47.36448	38.03237
TGARCH	-0.14580	0.16563	0.43412	21	42.84130	28.79639
SWARCH-L	-0.19103	0.28283	0.60511	12	24.31538	7.11071
HS (500 days)	0.11855	0.22221	0.25452	15	35.06016	13.16176
HS (1000 days)	0.11884	0.27441	0.20142	14	37.48319	10.99398
Adjusted HS (500 days)	0.09617	0.15559	0.43255	6	15.68104	0.18988
Adjusted HS (1000 days)	0.16566	0.19829	0.43404	5	11.96878	0.00000

Panel D. CAC40

	MRB	RMSRB	Correlation	Binary Loss Function	Quadratic Loss Function	LR(PF Test)
95%VaR						
EWMA	-0.01311	0.20403	0.34466	31	170.19216	1.41302
TGARCH	-0.05779	0.14269	0.37861	35	170.20397	3.76508
SWARCH-L	0.10484	0.33380	0.50413	16	124.33678	3.88827
HS (500 days)	-0.07911	0.17243	0.19004	43	250.01854	7.29855
HS (1000 days)	-0.07409	0.19250	0.10000	43	258.49763	11.33078
Adjusted HS (500 days)	0.04358	0.16129	0.37681	25	142.37146	0.00000
Adjusted HS (1000 days)	0.07567	0.23685	-0.21931	26	142.61332	0.04158
99%VaR						
EWMA	-0.16419	0.21753	0.34466	18	133.62785	20.45806
TGARCH	-0.20039	0.21643	0.37861	22	127.63642	31.78124
SWARCH-L	-0.05862	0.27293	0.50413	9	93.00723	2.61257
HS (500 days)	0.03070	0.17173	0.22891	11	119.68928	5.41909
HS (1000 days)	0.22108	0.34960	0.11190	8	96.87624	1.53828
Adjusted HS (500 days)	0.04195	0.11072	0.36937	3	83.68646	0.94312
Adjusted HS (1000 days)	0.12947	0.16603	0.37775	1	73.11995	4.81336

Figure 1 CUSUM of Squares Test



Note: The statistic of CUSUM of squares is used to test for the stability of the variance process. As with the CUSUM of squares test, movement outside the critical lines is suggestive of parameter or variance instability.