This article was downloaded by: [National Taiwan University]

On: 28 August 2008

Access details: Access Details: [subscription number 788856278]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

Applied Economics Letters

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t713684190>

Estimating value-at-risk via Markov switching ARCH models - an empirical study on stock index returns

Ming-Yuan Leon Li ^a; Hsiou-wei William Lin ^b

^a Department of Banking and Finance, National Chi Nan University, Taiwan ^b Department of International Business, National Chi Nan University, Taiwan

Online Publication Date: 15 September 2004

To cite this Article Li, Ming-Yuan Leon and Lin, Hsiou-wei William(2004)'Estimating value-at-risk via Markov switching ARCH models - an empirical study on stock index returns',Applied Economics Letters,11:11,679 — 691

To link to this Article: DOI: 10.1080/1350485042000236539

URL: <http://dx.doi.org/10.1080/1350485042000236539>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Estimating value-at-risk via Markov switching ARCH models – an empirical study on stock index returns

MING-YUAN LEON LI* and HSIOU-WEI WILLIAM LIN‡

Department of Banking and Finance, and ‡ Department of International Business,
National Chi Nan University, Taiwan

This paper estimates the Value-at-Risk (VaR) on returns of stock market indexes including Dow Jones, Nikkei, Frankfurt Commerzbank index, and FTSE via Markov Switching ARCH (SWARCH) models. It is conjectured that structural changes contribute to non-normality in stock return distributions. SWARCH models, which admit parameters based on various states to control structural changes in the estimating periods, may thus help mitigate kurtosis, tail-fatness and skewness problems in estimating VaR. Significant kurtosis and skewness in return distributions of Dow Jones, Nikkei, FCI and FTSE and significant tail-fatness (tail-thinness) in the 1% (5%) region critical probability are documented. Moreover, it is shown that the more generalized SWARCH outshines both ARCH and GARCH in capturing non-normalities with respect to both in- and out-sample VaR violation rate tests.

I. INTRODUCTION

The Markov switching ARCH (hereafter, SWARCH) model by Hamilton and Susmel (1994) is adopted to mitigate Kurtosis, tail-fatness and skewness in estimating Value-at-Risk (hereafter, VaR) for portfolios of stock market indexes. Specifically, a two-regime SWARCH setting is used for both mean and variance parameters of Dow Jones, Nikkei, Frankfurt Commerzbank Index (hereafter, FCI), and FTSE 100 returns. The findings support the notion that SWARCH models help capture fat tails, leptokurtosis and skewness in VaR estimation.

Most practitioners turned their attention to VaR no earlier than the mid-1990s. The Basle Committee on Banking Supervision of Bank for International Settlements (BIS) presented Amendment to Capital Accord to Incorporate Market Risks in January 1996. Many financial institutions now diversify their operations out of their original

businesses and actively trade capital market securities, foreign currencies and derivative instruments. These institutions' risks keep increasing because of increased return volatility, new competition, and deregulation. In April 1996, the Basle Committee further endorsed the use of value-at-risk measures for banks' capital adequacy ratios. By the end of 1997, G10 bank regulators took market risk into account for determining the risk-based assets.¹ Financial institutions have two alternatives. The first alternative, or the standardized risk measurement proposal, appears to be *ad hoc* and subject to strong assumption of homogeneity among the banks. The second alternative, also named the Value-at-Risk approach, in contrast, allows the users to gain accuracy in estimating market risks via tailored VaR methods.²

VaR is also at the centre of the recent interest in the risk management field.³ It serves to measure the level of market risks and accordingly, the capital adequacy ratios.

* Corresponding author. e-mail: lmy@ncnu.edu.tw

¹ The G10 group includes Belgium, Canada, France, Germany, Italy, Japan, Netherlands, Sweden, Switzerland, UK, USA and Luxembourg.

² This statement is especially descriptive for banks with large trading accounts.

³ One of the most frequently adopted VaR models is the *RiskMetrics* of J. P. Morgan Financial Service Co. Incorporated. *RiskMetrics*, which is publicly available on <http://www.jpmorgan.com>, is a representative parametric model.

Estrella *et al.* (1994), Kupiec (1995), Jackson *et al.* (1997) explore banks' capital adequacy ratios via the VaR approach. Duffie and Pan (1997) further suggest that default risk, credit risk, operation risks and liquidity risk could be measured via the algorithm of VaR.

Many researchers adopt non-parametric historical simulation and bootstrapping to estimate VaR. These approaches, nevertheless, depend on the prior period data. Take earthquake prediction as an example. The historical simulation would assign zero tail probability to, say, seven-scaled events if there had been no earthquakes that were equally or more severe in the estimation period. In contrast, parametric measures would recognize the seven-scaled threats when there had happened many significant, say, five-scaled, earthquakes. From such perspective, the parametric approaches may be more suitable than non-parametric ones. Prior investment risk studies, nevertheless, document non-normal histograms for financial asset returns. Typical non-normality properties include leptokurtosis, tail-fatness and skewness.

One potential explanation to the non-normality properties is the structural changes during the estimation period. As illustrated in Fig. 1, suppose an observation may either be drawn from Distribution 1 (as shown in Fig. 1(a)) with probability p or from Distribution 2 (as shown in Fig. 1(b)) with probability $1-p$. Let $X_{11}, X_{12}, X_{13}, \dots$, be the observations drawn from Distribution 1, which has a standard deviation of σ_1 , and $X_{21}, X_{22}, X_{23}, \dots$, be the observations from Distribution 2, which has a standard deviation of σ_2 . Statistically, when there exists a structural change from Distribution 1 to Distribution 2 during the past 1000 trading days (as shown in Fig. 1(c)), a researcher who equally weights these 1000 observations is likely to perceive the occurrences as being with Kurtosis, tail-fatness (as shown in the packed Fig. 1(d)) or even skewness.

Accordingly, it is proposed to control persistence of the volatility state via Markov-chain setting. The notion of volatility persistence is consistent with a large probability for a stage with low (high) volatility in returns to be followed by a low (high) volatility stage. Take sunspot and earthquake activities as examples. They both have persistent energy absorption and dissipation processes. A larger set of historical data may lead to better insights as to predicting the number of sunspots. Likewise, earthquake track records show that the mechanisms of energy absorption and energy dissipation are both cyclical. Also, crush motions in the neighbouring periods are serially

correlated. Thus scientifically it is foreseeable that an earthquake would trigger after-shocks in the subsequent periods.

For real financial world practices, it may be possible to mitigate the problem if unequal weights are assigned to the observations with differential ages. If there exists persistence, when most recent stock returns appear to be drawn from the more volatile distribution, it may be more (less) likely that the subsequent daily returns would also be drawn from the more (less) volatile distribution.

The settings with constant variance have long been questioned by the researchers. Moreover, both ARCH and GARCH are with linear volatility settings and may thus be less effective for studies sensitive to the changes in stock return volatility than SWARCH models,⁴ which incorporates Markov-switching and ARCH models, self-sufficiently partition different regimes and verify the point of structural changes based on historical data to partition the sample period as various phases.⁵

This paper presents the following results. First, the superiority of SWARCH over the linear, ARCH and GARCH models in capturing kurtosis, skewness, tail-fatness in the 1% critical region, as well as tail-thinness in the 5% region is demonstrated. Second, it is found that SWARCH outshines both ARCH and GARCH in out-sample violation rate tests.

The second section presents the definition for VaR and the competing models. Section 3 presents the empirical results. Section IV extends SWARCH models to estimate a multivariate system. Finally, Section V concludes the study and discusses potential future areas of work.

II. DEFINITION OF VAR AND SPECIFICATIONS FOR RETURN DISTRIBUTIONS

Set the critical probability to be α , then VaR_α , the absolute value at risk, is the expected maximum (worst) loss associated with a portfolios over a target time horizon within the confidence interval. Namely, VaR_α is the lower bound for equation $\int_{\text{VaR}_\alpha}^{+\infty} f(R_P) dR_P = 1 - \alpha$, where R_P may be the daily returns of a portfolio with a probability distribution function $f(R_P)$. Taking the firms' opportunity costs into account, one can further derive the relative value at risk, which is defined as the difference between the absolute VaR and the expected returns.

⁴ For instance, Hendricks (1994), Beder (1995), Simons (1996), Fong and Vasicek (1997) investigate the strengths and weaknesses of various VaR measures including parametric, historical simulation and Monte Carlo simulation methods. Jorion (1997) employs the student t -distribution in estimating VaR. Venkataraman (1997) adopts the mixture of normal distributions and the quasi-Bayesian estimation techniques for measuring the VaRs for a sample of eight exchange rates.

⁵ This study is also in contrast with Venkataraman (1997), who uses the binominal probability distributions for jumps from one distribution to another. Both mixing normal and MS incorporate two or more distributions to form a new distribution. Nevertheless, if the *perceived* non-normality is due to the above-mentioned structural change events, MS models should outperform mixing-normal models. Specifically, if there exists persistence and if period t returns are drawn from the first distribution, then the probability that the period $t+1$ return is drawn from the same distribution should be greater than p .

Fig. 1. PDF of Dow Jones index return shocks via the linear, ARCH, GARCH and SWARCH models

The key to VaR estimation is to estimate the return distribution based on the existing information set. Let R_t denote the returns for a specific portfolio at date t . There exist the following three competing specifications.

- (1) **Linear models:** The linear models define R_t as

$$R_t = u + \sigma \cdot e_t$$

In this setting, u and σ represent mean and the standard deviation of R_t , respectively. Also, e_t , which denotes the return shock, follows a standard Gaussian distribution.

- (2) **ARCH and GARCH models:** GARCH(p, q) models introduced by Bollerslev (1986) is as follows:

$$R_t = u + \sigma_t e_t \quad e_t \text{i.i.d.} \sim ND(0,1)$$

$$\varepsilon_t = \sigma_t e_t$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^q a_i \varepsilon_{t-i}^2 + \sum_{i=1}^p b_i \sigma_{t-i}^2$$

where e_t follows a Gaussian distribution with a standard deviation of unity. ARCH(q) models, which are proposed by Engle (1982) can be viewed as a special case of GARCH (p, q) with $p=0$. In the GARCH settings, the conditional variance may be affected by prior-period conditional variance and error sum of squares. λ , the volatility persistence measure, can be expressed as:

$$\lambda = (a_1 + a_2 + \dots + a_q + b_1 + b_2 + b_p)$$

- (3) **SWARCH models:** Hamilton and Susmel (1994) first establish the following SWARCH settings:

$$R_t = u_{s_t} + \varepsilon_t$$

$$\varepsilon_t = (g_{s_t} w_t)^{1/2}$$

$$w_t = (h_t e_t)^{1/2}$$

$$h_t = a_0 + a_1 w_{t-1}^2 + a_2 w_{t-2}^2 + \dots + a_q w_{t-q}^2$$

where e_t follows a Gaussian distribution with a standard deviation of unity and s_t is an unobservable state variable with possible outcomes of $1, 2, \dots, K$.

Further assume that s_t follows a one-order Markov process:

$$\begin{aligned} P(s_t = j | s_{t-1} = i, s_{t-2} = k, \dots, y_{t-1}, y_{t-2}, \dots) \\ = P(s_t = j | s_{t-1} = i) = p_{ij} \end{aligned}$$

Also define the transition probability matrix

$$P = \begin{bmatrix} P_{11} & P_{21} & \dots & P_{k1} \\ P_{12} & P_{22} & \dots & P_{k2} \\ \vdots & \vdots & \dots & \vdots \\ P_{1K} & P_{2K} & \dots & P_{kK} \end{bmatrix}$$

of which the sum of the elements in each and every row equal to one.

In SWARCH settings, w_t is a typical ARCH(q) process. Furthermore, ε_t is equal to w_t in ARCH(q) multiplied by a constant $(g_n)^{1/2}$ when $s_t=n$, where $n=1, 2, 3, \dots$, etc. Without loss of generality, g_1 , the coefficient for regime 1, is set to be unity and $g_i > 1$, $i=2, 3, \dots, k$ for the other regimes. Specifically, e_t is equal to w_t in an underlying fundamental ARCH(q) process multiplied by the square root of regime switching coefficient g_{s_t} . In a special case with $g_1 = g_2 = \dots = g_k = 1$, error term e_t in the model would follow an ARCH(q) process.⁶ In a two-regime setting, $s_t=1$ ($s_t=2$) may be set for the regime with low (high) return volatility. The transition probabilities can be presented as:⁷

$$p(s_t = 1 | s_{t-1} = 1) = p_{11}$$

$$p(s_t = 2 | s_{t-1} = 1) = p_{12}$$

$$p(s_t = 2 | s_{t-1} = 2) = p_{22}$$

$$p(s_t = 1 | s_{t-1} = 2) = p_{21}$$

where $p_{11} + p_{12} = p_{21} + p_{22} = 1$

In estimating regime variable s_t and regime probabilities at time t , conditional probability $p(s_t | y_{t+r}, y_{t+r-1}, \dots)$ may be tailored for different applications with simply substituting a different measure of r . Specifically, $p(s_t | y_{t+r}, y_{t+r-1}, \dots)$ is (1) a filtering probability when $r=0$, (2) an *ex ante* predicting probability when $r < 0$, or (3) an *ex post* smoothing probability when $r > 0$.⁸

III. EMPIRICAL ANALYSES

This study alternately adopts ARCH, GARCH and SWARCH to control non-normality properties of Dow Jones, Nikkei, FCI and FTSE index returns. For Dow Jones, Nikkei and FCI indexes, the sample period is between 7 January 1980 and 26 February 1999, whereas the sample period for FTSE only dated back to the beginning of 1984. It is demonstrated that Markov-switching setting effectively captures the structural changes in return volatility and mitigates downward biases in the loss

⁶ SWARCH incorporates MS and ARCH models, with the former setting filtering out return volatility, and ARCH controlling residual return volatilities. Specifically, MS models incorporate the discrete state variables to control the structural changes in the test period and mitigate persistence problems for estimating volatility in ARCH and GARCH. MS models therefore help capture kurtosis, skewness and tail-fatness in VaR estimation.

⁷ Note that, in contrast with Hamilton and Susmel (1994), who specify three regimes for return volatility, this study selects a more simplified two-regime setting. With limited number of prior observations, two- instead of three-regime settings may be sufficient.

⁸ Specifically, when the information set includes signals dated up to time t , the regime probability is referred to as a *filtering probability*, $p(s_t | y_t, y_{t-1}, \dots)$. On the other hand, when the overall sample period information set is used to estimate the state at t , the probability may be referred to a *smoothing probability*, $p(s_t | y_T, y_{T-1}, \dots)$. In contrast, a *predicting probability*, $p(s_t | y_{t-1}, y_{t-2}, \dots)$, is the regime probability for *ex ante* estimation, with the information set including signals dated up to period $t-1$.

Table 1. Estimates of skewness, kurtosis as well as 1%, 2.5%, and 5% critical values for the index returns shocks

Coefficients	Dow Jones	FCI	FTSE	Nikkei
Skewness coefficients ($N=0$)	-2.26	-0.47	-0.53	0.17
Kurtosis coefficients ($N=3$)	58.21	14.17	22.92	18.03
1% left-tailed critical value ($N=-2.33$)	-2.43	-2.46	-2.49	-2.78
2.5% left-tailed critical value ($N=-1.96$)	-1.90	-1.69	-1.87	-2.10
5% left-tailed critical value ($N=-1.65$)	-1.45	-1.26	-1.46	-1.55
1% right-tailed critical value ($N=2.33$)	2.43	2.24	2.32	2.82
2.5% right-tailed critical value ($N=1.96$)	1.92	1.48	1.75	1.97
5% right-tailed critical value ($N=1.65$)	1.44	1.15	1.36	1.42
Number of observations	4,838	4,758	3,801	5,045

Notes:

1. The return shock e_t is defined as $(R_t - \mu)/\sigma$, where (1) R_t denotes the t -th period stock index return and (2) μ and σ denote mean and standard deviation, respectively.
2. Benchmark N is the corresponding measure for a standard normal distribution.
3. This table shows that substantial kurtosis and skewness exist for every return distribution in the sample.
4. In both 1% and 2.5% regions, Nikkei return shock distribution has fatter tails on both sides as opposed to standard normal distributions. Tail fatness exists in the 1% regions for Dow Jones, FCI and FTSE index returns.

estimates from the prevalent linear, ARCH and GARCH models due to tail-fatness.⁹

Non-normality problems in the index return distributions

Table 1 presents descriptive statistics for kurtosis and skewness coefficients as well as the critical values in 1%, 2.5% and 5% tailed regions for the return shocks, indicating the significance of non-normality properties in the sample index returns. The significance of tail-fatness and tail-thinness in the distributions is further explored and it is documented that the absolute left-tailed 1% critical values for all the index returns are greater than 2.33 (the absolute 1% critical value of standard normal distribution). Conversely, the absolute left 5% critical values are less than 1.645 (the absolute 5% critical value of standard normal distribution). Namely, tail-fatness (tail-thinness) exists in the 1% (5%) critical region for all the index returns. Moreover, less pronounced but significant tail fatness exists in the left 2.5% region of the return shock distribution of Nikkei. The absolute left-tailed 2.5% critical value for Nikkei returns are greater than 1.96.

The performance in mitigating non-normality problems

The number of orders of prior-period error sum of squares in ARCH and SWARCH is set to be two. Moreover, the number of prior-period conditional variance and the number of error sum of squares in GARCH is set to

be one.¹⁰ Then 30 sets of initial parameter values are randomly generated and derive the maximum likelihood (ML) function value for each set of initial values. The mapped converged measure of the greatest ML function value then serves to estimate the SWARCH parameters.¹¹ Table 2 provides descriptive evidence on the three competing models for the various indexes. It shows that, with criteria including the maximum value of likelihood function, AIC and Schwarz value, SWARCH appears to outperform ARCH and GARCH.¹² Furthermore, consistent with the notion of two distinct volatility regimes of the index returns, all the g_2 estimates are significantly greater than one.¹³ Table 2 also shows that the estimates of u_1 (u_2) are significantly positive (insignificant) for all the indexes, lending supports to the notion of lower (greater) return volatility coinciding greater (lower) return means during the test period.

Table 3 presents the extent to which the competing models capture non-normalities for the index returns. Panel A demonstrates that the return shock kurtosis of SWARCH is significantly closer to three than the measures for the linear, ARCH and GARCH models. Panel B shows that the absolute skewness coefficients with respect to SWARCH are less than 0.1 and are significantly closer to zero than the measures for ARCH and GARCH. Consistently, Panel C shows that SWARCH outperforms the competing models in capturing (1) tail-fatness in the left 1% critical region and (2) tail-thinness in the left 2.5%

⁹ See Simons (1996) and Ho and Lin (1999).
¹⁰ The parameter estimates for both specifications of higher-ordered prior-period conditional variance and specifications of higher-ordered error sum of squares are insignificantly different from zero.
¹¹ OPTIMUM, a GAUSS package program, and the built-in Boyden, Fletcher, Goldfarb, and Shanno (BFGS) algebra are used to compute the negative minimum likelihood function. BFGS algebra is effective for deriving the maximum value of non-linear likelihood functions. See Luenberger (1984).
¹² Refer to Schwarz (1978) for Schwarz value and Akaike (1976) for AIC.
¹³ For example, the estimate of g_2 for Dow Jones is 5.292, indicating that the standard deviation of regime 2 returns is 5.292 times the standard deviation of regime 1 returns.

Table 2. The estimates of parameters of three non-linear models in capturing non-normalities

Panel A ARCH model (number of parameters = 4)											
	u	a_0	a_1	a_2	Log-Lik.	AIC	Schwarz				
Dow Jones	0.07 (0.013)	0.687 (0.019)	0.158 (0.018)	0.132 (0.018)	−6597.3	−6601.3	−6614.3				
FCI	0.07 (0.015)	0.84 (0.028)	0.267 (0.027)	0.186 (0.021)	−7270.0	−7274.0	−7287.0				
FTSE	0.068 (0.014)	0.568 (0.021)	0.145 (0.021)	0.214 (0.025)	−4984.3	−4988.3	−5000.8				
Nikkei	0.075 (0.012)	0.51 (0.02)	0.444 (0.031)	0.399 (0.033)	−7416.7	−7420.7	−7433.8				
Panel B GARCH model (number of parameters = 4)											
	u	a_0	a_1	b_1	Log-Lik.	AIC	Schwarz				
Dow Jones	0.070 (0.012)	0.022 (0.004)	0.079 (0.008)	0.901 (0.010)	−6397.2	−6401.2	−6414.2				
FCI	0.209 (0.015)	0.349 (0.032)	0.608 (0.052)	0.456 (0.021)	−7844.5	−7848.5	−7861.4				
FTSE	0.075 (0.013)	0.046 (0.008)	0.140 (0.014)	0.819 (0.018)	−4951.6	−4955.6	−4968.1				
Nikkei	0.074 (0.012)	0.056 (0.006)	0.244 (0.019)	0.750 (0.016)	−7128.4	−7132.4	−7145.5				
Panel C SWARCH model (number of parameters = 8)											
	p_{22}	p_{11}	u_1	u_2	a_0	a_1	a_2	g_2	Log-Lik.	AIC	Schwarz
Dow Jones	0.919 (0.024)	0.987 (0.003)	0.072 (0.010)	−0.005 (0.050)	0.531 (0.025)	0.069 (0.011)	0.039 (0.010)	5.292 (0.638)	−6339.3*	−6347.3*	−6373.2*
FCI	0.843 (0.031)	0.988 (0.002)	0.081 (0.014)	−0.263 (0.174)	0.587 (0.025)	0.142 (0.023)	0.135 (0.028)	21.648 (2.855)	−6888.5*	−6896.5*	−6922.4*
FTSE	0.958 (0.015)	0.995 (0.002)	0.065 (0.014)	0.043 (0.033)	0.506 (0.021)	0.062 (0.015)	0.123 (0.022)	5.446 (0.645)	−4859.4*	−4867.4*	−4892.4*
Nikkei	0.982 (0.004)	0.988 (0.003)	0.072 (0.011)	−0.023 (0.062)	0.236 (0.011)	0.189 (0.023)	0.151 (0.021)	8.835 (0.479)	−6781.0*	−6789.0*	−6815.1*

Notes:

1. In the parentheses, the estimates of standard deviation are presented.
2. According to Akaike (1974), $AIC = ML \text{ function value} - N$, where N is the number of parameters.
3. According to Schwarz (1978), $Schwarz \text{ value} = ML \text{ function value} - (N/2) \times \ln(T)$, where T is the number of observations.
- 4.* denotes the maximum value among Panels A, B and C.

and 5% regions. Furthermore, Panel D shows that Jarque–Bera (JB) normality test statistics, a criterion for both skewness and kurtosis, with respect to SWARCH are significantly smaller and closer to zero than those for the alternative models. The findings support the notion that SWARCH provides the closest approximation of normal return shocks.

Figure 2 present the illustrative charts of probability density functions (hereafter PDFs) estimated by the competing models, as opposed to the standard normal setting, for Dow Jones return shocks. Figure 2(a) shows that both kurtosis and tail-fatness in the 1% region are significant for the linear model. Figs (b), (c) and (d) document that, with (a) as the benchmark, both kurtosis and 1% tail-fatness for ARCH, GARCH and SWARCH

estimates are significantly less pronounced. The results suggest that the three non-linear alternatives help capture non-normalities. Nevertheless, ARCH and GARCH still generate pronounced kurtosis and tail-fatness in extreme tailed regions. In contrast, the return shocks estimated by SWARCH closely conform to standard normal distribution, suggesting that SWARCH settings with discrete state variable help filter the structural changes and mitigate non-Gaussian problems.

Out-sample violation rate tests

Horse races were conducted for linear, ARCH, GARCH and SWARCH models on Dow Jones, FCI, FTSE and Nikkei returns with 1000-day windows in the rolling

Table 3. The performance of various models in capturing non-normalities

Index return series	Linear	ARCH	GARCH	SWARCH
Panel A Kurtosis coefficients (the kurtosis coefficient for a standard distribution is 3)				
Dow Jones	58.21	9.05	9.67	3.44*
FCI	14.17	22.24	23.84	4.69*
FTSE	22.92	8.73	12.30	3.10*
Nikkei	18.03	13.68	30.69	3.81*
Panel B Skewness coefficients (the skewness coefficient for a standard normal distribution is 0)				
Dow Jones	-2.26	-0.43	-0.57	-0.08*
FCI	-0.47	-0.53	-0.62	0.09*
FTSE	-0.53	-0.31	-0.25	-0.08*
Nikkei	0.17	-0.50	-1.19	-0.02*
Panel C Tail-fatness and tail-thinness coefficients				
1% critical value ($N = -2.33$)				
Dow Jones	-2.44	-2.55	-2.51	-2.22*
FCI	-2.46	-2.49	-2.36	-2.31*
FTSE	-2.49	-2.57	-2.53	-2.37*
Nikkei	-2.78	-2.79	-2.51	-2.26*
2.5% critical value ($N = -1.96$)				
Dow Jones	-1.90	-2.01	-1.99	-1.96*
FCI	-1.69	-1.87	-1.81	-2.00*
FTSE	-1.87	-1.99	-1.95	-1.96*
Nikkei	-2.10	-2.03	-2.05	-1.93*
5% critical value ($N = -2.33$)				
Dow Jones	-1.45	-1.58	-1.57	-1.64*
FCI	-1.26	-1.43	-1.47	-1.66*
FTSE	-1.46	-1.59	-1.62	-1.69*
Nikkei	-1.55	-1.62	-1.63	-1.65*
Panel D Jarque-Bera statistics				
Dow Jones	617279.00	7519.35	10256.17	32.07*
FCI	94543.10	114055.40	115509.20	153.33*
FTSE	33919.96	3746.00	8320.54	19.20*
Nikkei	47386.61	24144.70	170511.40	129.23*

Notes:

1. Jarque-Bera = $N [S^2/6 + (k-3)^2/24]$, where N is the numbers of observation, and S and K represent the skewness and kurtosis coefficients respectively.
2. * indicates the measure closest to the value for standard normal distribution in the row.
3. For a standard normal distribution, Jarque-Bera statistic is 0.

estimation process.¹⁴ The research design begins with collecting the 1000 pre-VaR daily returns, $\{R_{i-t}\}_{i=1}^{1,000}$, for each date t . Then 10 000 random numbers are generated to simulate the distribution of index return R_t via Monte Carlo algorithm. Furthermore VaR_t is estimated for 1%, 2.5% and 5% left-tailed regions. For each critical value, the number of observations with return $R_t < VaR_t$ is taken as the number of *left-tailed* violations.¹⁵ An *ex ante* predicting probabilities $P(s_t|y_{t-1}, y_{t-2}, \dots)$ are adopted and two normal distributions weighed with different means and variances. Also the *VaR* algorithm is applied to the right tails. Likewise, critical values are estimated for *right* 1%, 2.5% and 5% regions. For

each critical value, the number of observations with $R_t > VaR_t$ is taken as the number of *right-tailed* violations. Then the number of violations is divided by the number of out-sample observations to derive the violation rates.

Consider the Dow Jones series as an example. There are 4838 trading days during the sample period. For the tests with 1000 prior-trading-day estimation window and one-day as the order of the lagged term, there are 3837 out-sample observations of violation rates. Figure 3(a-d) present Dow Jones returns, the estimates of g_2 for regime 2, the estimated mean returns for regimes 1 and 2 (u_1 and u_2) and the predicting probabilities with respect to regime 2,

¹⁴ The Basle Committee proposes that the learning window should at least include 250 pre-VaR trading days. In the rolling estimation process used, one more observation are introduced for each time point. To facilitate convergence in the process, the estimate for each period is used as the initial value in the non-linear estimation for the immediately subsequent period.

¹⁵ It is a common practice of estimating one-holding-day VaRs.

Fig. 2. Daily Dow Jones index returns and the parameter estimates of SWARCH model with 1,000 prior-trading-day Windows

Fig. 3. The predicting VaR via SWARCH model with 1,000-Prior-trading-day windows for various confidence intervals

Fig. 4. The Four Types of Correlation Coefficients for the Portfolio established by Dow Jones and FTSE Stock Assets with Equal Weight and the Predicting VaR with 1,000-Prior-trading-day Windows via Linear and SWARCH model

respectively. Since the estimates of g_2 are significantly greater than one, regime 1 (regime 2) is label as the high (low) volatility regime.¹⁶

Figure 4(a–d) present the VaRs estimated by SWARCH for 95%, 97.5% and 99% confidence intervals, respectively. To include both regions of excessive losses and gains, the critical values on both tails of the distributions are estimated. Namely, the measures of violations used include *left-tailed*, *right-tailed* and the aggregate of *left- and right-tailed* violations.

Table 4 presents the violation rates for the linear, ARCH, GARCH and SWARCH models. A log-likelihood test is adopted for the null hypothesis of violation rate equals the critical probability.¹⁷ If the model fails to reject (rejects) the null hypothesis, then it is concluded that the model is accurate (inaccurate).

Panel A of Table 4 presents the violation rates associated with the 5% critical regions. As to the aggregate violations, SWARCH is accurate for all the indexes, whereas the other models are accurate for some but inaccurate for some other indexes. Furthermore, because of high corresponding *left-tailed* violation rates, the null hypotheses that the linear model is accurate for FTSE

and Nikkei returns are rejected (significant at 1% level). Moreover, ARCH and GARCH are inaccurate left-tailed models for Nikkei but accurate left-tailed models for the other index returns. In contrast, the null hypotheses that SWARCH is an accurate *left-tailed* model in estimating VaRs cannot be rejected for all the index returns. With respect to the right-tailed regions, each specification appears to be accurate for some indexes but inaccurate for some other indexes.

Panel B presents the violation rates associated with the 2.5% regions. Consistently, the null hypothesis that SWARCH is accurate with respect to all the *left-tailed*, *right-tailed* and aggregate violation tests cannot be rejected. In contrast, the linear and ARCH models are inaccurate in estimating the *left-tailed* 2.5% VaRs for Dow Jones, FCI and Nikkei, whereas GARCH is inaccurate for Nikkei.

Consistent with the notion of more pronounced tail-fatness towards extremes of the returns distributions, the numbers of both * and ** in Panel C are larger than the measures in Panels A and B. Namely, there are significantly more (less) rejects of the null hypothesis for the 1% critical regions (the 2.5% and 5% regions.) In Panel C, the *left-tailed*, the *right-tailed* and the aggregate violation rates

¹⁶ In contrast, there exist mixed results as to the order of sample means with respect to regimes 1 and 2.

¹⁷ Lopez and Walter (2001) and Engle and Manginelli (1999) may be referred to for other statistical tests for VaR estimates.

Table 4. Violation rates for the linear, ARCH, GARCH and SWARCH models for the sample stock index returns

Stock Index Returns	Linear model			ARCH model			GARCH model			SWARCH model		
	Left	Right	Mean	Left	Right	Mean	Left	Right	Mean	Left	Right	Mean
Panel A Critical probability = 5%												
Dow Jones	4.40% (2.979)	4.72% (0.658)	4.56% (1.603)	4.85% (0.186)	4.74% (0.536)	4.79% (0.338)	4.48% (2.226)	4.38% (3.239)	4.43% (2.707)	4.35% (3.537)	4.72% (0.658)	4.53% (1.802)
FCI	5.24% (0.462)	4.23% (4.909)*	4.74% (0.553)	5.32% (0.804)	3.73% (14.045)**	4.52% (1.852)	4.79% (0.354)	3.25% (27.628)**	4.02% (8.154)**	5.59% (2.653)	4.26% (4.566)*	4.93% (0.046)
FTSE	3.86% (8.353)**	3.82% (8.906)**	3.84% (8.627)**	4.61% (0.942)	3.71% (10.684)**	4.16% (4.411)*	4.53% (1.320)	3.07% (25.314)**	3.80% (9.190)**	5.04% (0.008)	3.75% (10.045)**	4.39% (2.262)
Nikkei	5.71% (4.122)*	5.04% (0.016)	5.38% (1.183)	5.99% (7.779)**	4.48% (2.422)	5.23% (0.444)	5.76% (4.718)*	3.78% (13.709)**	4.77% (0.407)	5.00% (0.000)	4.15% (6.445)*	4.57% (1.583)
Panel B Critical probability = 2.5%												
Dow Jones	3.07% (4.860)*	2.61% (0.175)	2.84% (1.752)	3.20% (7.221)**	2.61% (0.177)	2.91% (2.475)	2.79% (1.271)	2.53% (0.013)	2.66% (0.390)	2.42% (0.091)	2.42% (0.091)	2.42% (0.091)
FCI	3.83% (23.605)**	2.98% (3.364)	3.41% (11.409)**	3.54% (14.776)**	2.32% (0.540)	2.93% (2.668)	2.93% (2.668)	1.76% (9.502)**	2.34% (0.395)	3.27% (3.427)	2.24% (1.115)	2.75% (0.072)
FTSE	2.50% (0.000)	2.21% (0.981)	2.36% (0.242)	2.86% (1.394)	2.43% (0.061)	2.64% (0.227)	2.18% (1.246)	1.61% (10.482)**	1.89% (4.629)*	2.64% (0.230)	2.00% (3.080)	2.32% (0.375)
Nikkei	3.66% (19.542)**	3.21% (7.768)**	3.44% (13.051)**	3.93% (29.043)**	2.67% (0.473)	3.30% (9.691)**	3.58% (17.271)**	2.05% (3.536)	2.82% (1.622)	2.89% (2.443)	2.30% (0.684)	2.60% (0.152)
Panel C Critical probability = 1%												
Dow Jones	1.85% (22.408)**	1.69% (15.452)**	1.77% (18.795)**	1.98% (29.019)**	1.51% (8.779)**	1.75% (17.665)**	1.77% (18.811)**	1.25% (2.266)	1.51% (8.799)**	1.22% (1.834)	1.25% (2.266)	1.24% (2.045)
FCI	2.58% (66.105)**	1.94% (26.460)**	2.26% (44.543)**	2.39% (53.103)**	1.30% (3.199)	1.85% (21.899)**	1.78% (18.875)**	1.01% (0.005)	1.40% (5.326)*	1.65% (13.416)**	1.01% (0.005)	1.33% (3.763)
FTSE	1.68% (10.803)**	1.32% (2.648)	1.50% (6.120)*	1.71% (11.874)**	1.46% (5.324)*	1.59% (8.319)**	1.43% (4.577)*	0.86% (0.609)	1.14% (0.549)	1.46% (5.333)*	1.11% (0.314)	1.29% (2.118)
Nikkei	2.55% (68.423)**	2.08% (36.141)**	2.31% (51.292)**	2.10% (37.659)**	1.71% (16.816)**	1.90% (26.389)**	1.90% (26.389)**	1.16% (1.022)	1.53 (9.983)**	1.21% (1.714)	1.11% (0.501)	1.16% (1.022)

Notes:

- * and ** indicate significance at the 5% (critical value is 3.841) and 1% (critical value is 6.6349) levels (for one-tailed tests), respectively.
- The log-likelihood test statistics (presented in parentheses) $LR = 2\{\ln[(\alpha^*)^x(1 - \alpha^*)^{T-x}] - [\ln(\alpha^x(1 - \alpha)^{T-x})]\}$, where T is the number of observations for comparing return and critical value, x is the number of violations, α is the true underlying critical probability, the sample violation rate $\alpha^* = x/T$. The test statistics is an asymptotic χ^2 distribution with one degree of freedom.

via both linear and ARCH specifications are significantly greater than 1% for all the indexes. Moreover, GARCH significantly underestimates all the (*left-tailed*) 1% VaRs but is accurate for the *right-tailed* 1% region with respect to all the indexes. In contrast, SWARCH is an inaccurate left-tailed model for FCI and FTSE returns but an accurate left-tailed model for Dow Jones and Nikkei returns. Moreover, as to the right-tailed and the aggregate violations, SWARCH is accurate for all the indexes. To sum up, SWARCH is the most accurate measure for the left 1% region.

The results in Table 4 also indicate the difficulties in estimating the frequency of rare adversarial events and may add to the contemporary literature, which mixes the *right-* and the *left-tailed* violations and focuses on the aggregate measures.¹⁸

IV. COMPARISON TESTS ON THE HIGHER DIMENSIONAL SYSTEM

This section extends SWARCH models to estimate a multivariate system. The features of the higher dimensional system are that the correlations among various assets in a portfolio should be considered. Nevertheless, SWARCH process becomes complicated when the analysis is extended to estimating portfolio risks. Specifically, this paper sets two outcomes of the discrete state variable s_t to represent high and low volatility regimes and two orders of prior-period error squares. Thus it is necessary to consider $2^3 = 8$ possible states for any univariate index return at each date. Therefore, in a portfolio with four assets, it is necessary to consider $8^4 = 4096$ possible state combinations.

Accordingly, here a computationally simpler design is adopted. First, predicting probability $p(s_t|y_{t-1}, y_{t-2}, \dots)$ is used to identify the volatility regimes of each index return at each date. Specifically, if $p(s_t = 1|y_{t-1}, y_{t-2}, \dots) > 0.5$ (< 0.5), then it is concluded that R_t is at the low (high) volatility state. Second, the paired-up sets with equal weights are exhausted of the four major indexes and form six portfolios including (1) 50% Dow Jones and 50% FTSE, (2) 50% FTSE and 50% FCI, (3) 50% FCI and 50% Nikkei, (4) 50% Nikkei and 50% Dow Jones, (5) 50% FTSE and 50% Nikkei, and (6) 50% FCI and 50% Dow Jones. The return from investing can be written in any one of the six portfolios as

$$r_{PORT} = 0.5r_i + 0.5r_j, \quad i \text{ not equal to } j$$

With the setting of two distinct volatility regimes for each index, the returns of each portfolio may be drawn from

four possible combinations of volatility regimes. For instance, the equally weighted portfolio of Dow Jones and FTSE may be with four combinations of volatility states including (1) both indexes being at low volatility states (LOW-LOW), (2) Dow Jones being at low volatility states but FTSE being at high volatility states (LOW-HIGH), (3) Dow Jones being at high volatility states but FTSE being at low volatility states (HIGH-LOW), and (4) both indexes being at high volatility states (HIGH-HIGH). Third, 1000-prior-trading-day windows are used in the rolling estimation process and the correlation coefficients measured for each combination of volatility regimes of the two component indexes. Fourth, the parameter estimates are used to simulate the distribution of index return R_t via Monte Carlo algorithm and the violation rates calculated.

Figure 4(a) presents the four series of estimated correlation coefficients between Dow Jones and FTSE index returns. It is worth noting that (1) the four correlation coefficients appear to be significantly different, (2) the correlation coefficients during the LOW-LOW stages are substantially less volatile, and (3) conversely, the correlation coefficients during the HIGH-HIGH stages are substantially more volatile. The findings suggest that the index return correlations are more (less) stable in lower (higher) variance regimes.¹⁹ Figure 4(b,c) presents the returns of the portfolio of Dow Jones and FTSE and the corresponding 1% VaR estimates for via linear and SWARCH models, respectively.²⁰

As demonstrated in the previous section, the superiority of SWARCH to the competing models is most pronounced with respect to the 1% critical probabilities. Thus the VaR comparison tests are focused on the 1% tailed regions and the linear models adopted as the benchmark.²¹ Table 5 presents the estimates and the log-likelihood statistics of violation rates for various portfolios. Consistent with the notion of structural changes of the portfolio returns, the violation rates corresponding to the linear model for all six portfolios are substantially greater than 1%. In contrast, as to the left-tailed SWARCH estimates, the null hypothesis of 1% violation rates for all six portfolios cannot be rejected.

V. CONCLUSIONS AND EXTENSIONS

This paper serves as one of the first systematic studies on applying SWARCH to VaR estimation. Prior VaR studies most typically encounter non-normalities in return distributions. Consistently, it is shown that leptokurtosis,

¹⁸ Refer to Venkataraman (1997), who assumes symmetry of the two ends and aggregates the tailed region measures.

¹⁹ Similar results are found for the other five portfolios.

²⁰ VaRs with 1000-prior-trading-day windows are estimated.

²¹ According to the above-mentioned results, the linear, ARCH and GARCH models are inaccurate in estimating the 1% VaRs.

Table 5. Higher dimensional violation rate tests in estimating the 1% portfolio VaRs with linear and SWARCH models

Paired-up portfolios	SWARCH model			Linear model		
	Left tail	Right tail	Mean	Left tail	Right tail	mean
50% Dow Jones and 50% FTSE	1.143% (0.553)	0.571% (6.158)*	0.857% (0.608)	1.607% (8.802)**	1.250% (1.638)	1.428% (4.574)*
50% Dow Jones and 50% FCI	1.107% (0.313)	0.750% (1.935)	0.929% (0.146)	1.928% (19.156)**	1.392% (3.873)*	1.660% (10.277)**
50% Dow Jones and 50% Nikkei	0.964% (0.037)	1.143% (0.553)	1.054% (0.081)	1.714% (11.880)**	1.857% (16.583)**	1.785% (14.133)**
50% FTSE and 50% FCI	1.143% (0.553)	0.536%(7.326)**	0.8395% (0.771)	1.999% (21.877)**	1.392% (3.873)*	1.696% (11.335)**
50% FTSE and 50% Nikkei	0.929% (0.146)	1.179% (0.857)	1.054% (0.081)	1.749% (12.970)**	1.749% (12.970)**	1.749% (12.970)**
50% FCI and 50% Nikkei	0.821% (0.965)	0.893% (0.336)	0.857% (0.608)	1.964% (20.517)**	1.571% (7.856)**	1.767% (13.547)**

Note:

1. The stock indexes are paired up and form six equally weighted portfolios. For instance, the returns of the ‘50% Dow Jones and 50% FTSE’ portfolio is computed as $0.5 \times (\text{Dow Jones returns}) + 0.5 \times (\text{FTSE returns})$.
2. The portfolio VaRs are estimated via Monte Carlo simulation in the following way:
 - (a) Draw a random number, z_i , from a standard normal distribution.
 - (b) Multiply z_i by the standard deviations of asset i index return (σ_i) and plus the mean of asset i index return (u_i) to create the first index returns for the scenario, f_i : $f_i = u_i + \sigma_i z_i$, $z_i \sim N(0, 1)$.
 - (c) Multiply z_i by the estimate of correlation, $\rho_{i,j}$.
 - (d) Draw a second independent random number z_j .
 - (e) Multiply z_j by the rood of one minus the correlation squared.
 - (f) Add the two results together to create a random number y that has a standard deviation of one and correlation $\rho_{i,j}$ with f_i : $y = z_i \rho_{i,j} + z_j (1 - \rho_{i,j}^2)^{0.5}$, $z_j \sim N(0,1)$.
 - (g) Multiply y by the standard deviation of j asset return, σ_j , and plus the mean of j asset return (u_j) to create the second index returns for that scenario, f_j .
 - (h) Create the equally weighted portfolio return for that scenario, f : $f = 0.5 \times f_i + 0.5 \times f_j$
 - (i) Create 10 000 scenarios, the 99% VaR would be the one-hundred-worst result.
3. Refer to Table 4 for the other notations.

fat tails and significant skewness exist in the return distributions of the major index returns. Furthermore, tail fatness appears to be more significant near both extremes.

It is conjectured that the incorporations of historical data without controlling prior regime changes would contribute to the documented non-normality problems. Therefore, among the alternative non-linear VaR models, SWARCH is proposed over the competing specifications. The findings lend supports to the superiority of SWARCH models in mitigating non-normalities of financial returns. Especially as to the 1% tailed region, SWARCH outshines the other measures and may be the best alternative to financial institutions, for which regulators and the risk managers often focus on the 1% or more extreme left-tailed region. SWARCH models are also extended to higher-dimensional VaR estimation via a computationally simpler design. Consistently, SWARCH appears to be an accurate model in the 1% critical region.

Future areas of work may include exploring the extent other parametric alternatives such as t -distribution and GED distribution help mitigating the non-normalities of financial asset returns.²² Moreover, the performance of the extended ARCH or GARCH models such as GARCH accounting for the leverage effect, exponential GARCH (EGARCH) and other asymmetric GARCH specifications

in estimating financial return distributions or volatility dynamics may warrant further studies.

ACKNOWLEDGEMENTS

The Authors are grateful to Darrel Duffie and James D. Hamilton for their invaluable comments and suggestions.

REFERENCES

Akaike, H. (1974) A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, **9**, 716–23.

Beder, T. S. (1995) VAR: seductive but dangerous, *Financial Analysts Journal*, September, 12–24.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **28**, 307–27.

Duffie, D. and Pan, J. (1997) An overview of value at risk, *Journal of Derivatives*, **4**, 7–49.

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

Engle, R. F. and Manginelli, S. (1999) CAViaR: conditional value-at-risk by quantile regression, *NBER Working Paper No. 7341*.

²²See Bollerslev *et al.* (1994) for the details regarding these approaches.

- Estrella, A., Hendricks, D., Kambhu, J., Shin, S. and Walter, S. (1994) The price risk of options positions: measurement and capital requirements, *Federal Reserve Bank of New York Quarterly Review*, **19**(2), 27–43.
- Hamilton, J. D. (1991) A quasi-Bayesian approach to estimating parameters for mixtures of normal distributions, *Journal of Business and Economic Statistics*, **9**, 27–39.
- Hamilton, J. D. and Susmel, R. (1994) Autoregressive conditional heteroscedasticity and changes in regime, *Journal of Econometrics*, **64**, 307–33.
- Jackson, P., Maude, D. J. and Peerraudin, W. (1997) Bank capital and value at risk, *Journal of Derivatives*, **4**, 73–89.
- Jorion, P. (1997) *Value at Risk: the New Benchmark for Controlling Market Risk*, Irwin Professional, Chicago, IL.
- Morgan, J. P. (1995) *RiskMetrics*, www.jpmorgan.com.
- Kupiec, P. H. (1995) Techniques for verifying the accuracy of risk measurement models, *Journal of Derivatives*, **3**(2), 73–84.
- Lopez, J. A. and Walter, C. A. (2001) Evaluating covariance matrix forecasts in a value-at-risk framework, *Journal of Risk*, **3**(3), 69–98.
- Luenberger, D. G. (1984) *Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA.
- Schwarz, G. (1978) Estimating the dimension of a model, *Annals of Statistics*, **6**, 461–4.
- Simons, K. (1996) Value at risk – new approach to risk management, *New England Economic Review*, September/October, 3–13.
- Venkataraman, S. (1997) Value at risk for a mixture of normal distributions: the use of quasi-Bayesian estimation techniques, *Economic Perspectives*, March, 2–13.