Malaysian Journal of Mathematical Sciences **1(1)**: 83 - 102 (2007)

Modeling and Forecasting Volatility of the Malaysian and the Singaporean Stock Indices using Asymmetric GARCH Models and Non-normal Densities

[1]Abu Hassan Shaari Mohd Nor & [2]A. Shamiri
[1]Faculty of Economics and Business
[2]Faculty of Science and Technology, Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor, Malaysia
E-mail : ahshaari@yahoo.com

ABSTRACT

This paper examines and estimates the three GARCH(1,1) models (GARCH, EGARCH and GJR-GARCH) using daily price data. Two Asian stock indices KLCI and STI were studied using daily data over a 14-years period. The competing models include GARCH, EGARCH and GJR-GARCH using the Gaussian normal, Student-t and Generalized Error Distributions. The estimates showed that the forecasting performance of asymmetric GARCH Models (GJR-GARCH and EGARCH), especially when fat-tailed densities are taken into account in the conditional volatility, are better than symmetric GARCH. Moreover, it was found that the AR(1)-GJR model provides the best out-of-sample forecast for the Malaysian stock market, while AR(1)-EGARCH provides a better estimation for the Singaporean stock market.

Keywords: ARCH-Models, Asymmetry, Stock market indices and volatility modeling
JEL *classification*: G14;C13;C22.

INTRODUCTION

Traditional regression tools have shown their limitation in the modeling of high-frequency (weekly, daily or intra-daily) data. The assumption that only the mean response changes with covariates, while the variance remains constant over time has often revealed to be unrealistic in practice. This fact is particularly obvious in series of financial data where clusters of volatility can be detected visually. Indeed, it is now widely accepted that high frequency financial returns are heteroskedastic.

Modeling financial time series is not an easy task because they possess some special characteristics (Tsay, 2002). They often exhibit volatility clustering (i.e. large changes tend to be followed by large changes and small changes by small changes), leptokurtosis (i.e., the distribution of their returns is fat tailed) and leverage effect (i.e. changes in stock prices tend to be negatively correlated with changes in volatility which implies volatility is higher after negative than after positive shocks of the same magnitude). In order to capture the first two characteristics of financial time series, Engle (1982) proposed to model time-varying conditional variance with the Auto-Regressive Conditional Heteroskedasticity (ARCH) processes that use past disturbances to model the variance of the series. Early empirical evidence shows that high ARCH order has to be selected in order to catch the dynamics of the conditional variance. The Generalized ARCH (GARCH) model of (Bollerslev, 1986) is an answer to this issue. It is based on an infinite ARCH specification and it allows reducing the number of estimated parameters from ∞ to only 2. Both models allow taking the first two characteristics into account, but their distributions are symmetric and therefore fail to model the third stylized fact, namely the "leverage effect". To solve this problem, many nonlinear extensions of the GARCH model have been proposed. The widespread models include the Exponential GARCH (EGARCH) of (Nelson, 1991) and the so-called GJR (Glosten, Jagannathan, & Runkle, 1993).

Unfortunately, GARCH models often do not fully capture the thick tails property of high frequency financial time series. This has naturally led to the use of non-normal distributions to better model this excess kurtosis, such as Student-t distribution, generalized error distribution, Normal-Poisson, Normal-Lognormal and Bernoulli-Normal distributions. Liu and Brorsen (1995) introduced the use of an asymmetric stable density to capture the skewness property well. However, since the variance of such a distribution rarely exists, it is not popular in practice. Bollerslev (1986) introduced the Student-t distribution, which captures the kurtosis for heavy tailed data. Lambert & Laurent (2001) extended this to the GARCH model.

We have selected the Strait Times Index in Singapore (STI) and Kuala Lumpur Composite Index in Malaysia (KLCI) to investigate the behaviour of both markets. Pan et al. (1999) and Kim (2003) studied the GARCH effects to examine linkages between the U.S. and five Asian-Pacific stock markets (Australia, Hong Kong, Japan, Malaysia, and Singapore). Choudhry (2005) investigated the effects of the Asian financial crisis of 1997-1998 on the time-varying data of 10 firms each from Malaysia and Taiwan. A recent study by Cheong et al. (2007), investigates the long-memory behavior of the Malaysian Stock Exchange. While the research on evaluating each volatility model has been very versatile since the introduction of ARCH model by (Engle, 1982), there has been much less effort in comparing alternative density forecast models. However, another striking characteristic of high-frequency financial returns is that they are often characterized by fat-tailed distribution. In fact, the kurtosis of most asset returns is higher than three, which means that extreme values are observed more frequently that for the normal distribution. While the high kurtosis of the returns is a well-established fact, the situation is much more obscure with regard to the symmetry of the distribution. Many researchers have not observed anything special on this aspect, but others (Simkowitz & Beedles, 1980), (Kon, 1984) and (So, 1987) have highlighted the heavy tailed distribution. Mittnik and Paolella (2001) have shown that a fat-tailed distribution is required for modeling several daily exchange rate returns of East Asian currencies against the US dollar.

However, in this paper we demonstrate that this gap can be filled by a rigorous density forecast comparison methodology. We compare the performance of the GARCH, EGARCH and GJR-GARCH models and also introduce different densities (Normal, Student-t and GED).

EMPIRICAL METHODOLOGY

ARCH-Models

Over the past two decades, enormous effort has been devoted to modeling and forecasting the movement of stock returns and other financial time series. Seminal work in this area of research can be attributed to Engle (1982), who introduced the standard Autoregressive Conditional Heteroskedasticity (ARCH) model. Engle's process proposed to model time-varying conditional volatility using past innovations to estimate the variance of the series as follows:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \tag{1}$$

where ε_t denotes a discrete-time stochastic taking the form of $\varepsilon_t = z_t \sigma_t$ where $z_t \sim iid$ (0,1), and σ_t is the conditional standard deviation of return at time *t*, assuming that market returns follow AR(p) process as given below:

$$R_t = \varphi_0 + \sum_{i=1}^{q} \phi_i R_{t-i} + \varepsilon_t \tag{2}$$

GARCH

Further extension introduced by Bollerslev (1986) known as the Generalized ARCH (GARCH) model which suggest that the time-varying volatility process is a function of both past disturbances and past volatility. The GARCH model is an infinite order ARCH model generated by:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{3}$$

where α_0, α and β are non-negative constants. For the GARCH process to be defined, it is required that $\alpha > 0$.

EGARCH

The first asymmetric GARCH model that is looked at is the EGARCH model of Nelson (1991), which looks at the conditional variance and allows for the asymmetric relation between stock returns and volatility changes. Nelson indicates that by including an adjusting function *g(z)* in the conditional variance equation, it in turn becomes expressed as:

$$\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i g(z_{t-i}) + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-j}^2) \tag{4}$$

where $z_t = \varepsilon_t / \sigma_t$ is the standardized residual series.

The value of *g(zt)* is a function of both the magnitude and sign of *zt* and is expressed as:

$$g(z_t) = \underbrace{\theta z_t}_{sign\ \ effect} + \underbrace{\theta[|z_t| - E|z_t|]}_{magnitude\ \ effect} \tag{5}$$

Notice moreover that E|zt| depends on the assumption made on the unconditional density. This aspect is discussed further in Section 3. The EGARCH model differs from the standard GARCH model in two main aspects. First, it allows positive and negative shocks to have a different impact on volatility. Second, the EGARCH model allows large shocks to have a greater impact on volatility than the standard GARCH model.

GJR-GARCH

This model is proposed by Glosten, Jagannathan, and Runkle (1993). The generalized form is given by:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} (\alpha_i \varepsilon_{t-1}^2 + w_i S_{t-i}^- \varepsilon_{t-1}^2) + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \quad (6)$$

where S_t^- is a dummy variable.

In this model, it is assumed that the impact of ε_t^2 on the conditional variance σ_t^2 is different when ε_t is positive or negative. It is for this reason that the dummy variable S_t^- takes the value '0' (respectively '1') when ε is positive (negative). It is worth noting that the TGARCH model of (Zakoian, 1994) is very similar to GJR but TGARCH models the conditional standard deviation instead of the conditional variance.

DENSITIES ASSUMPTIONS

The GARCH models are estimated using the maximum likelihood (ML) methodology[1]. The logic of ML is to interpret the density as a function of the parameters set, conditional on a set of sample outcomes. This function is called the likelihood function.

Failure to capture the fat-tails property of high-frequency financial time series has led to the use of non-normal distributions to better model excessive third and fourth moments. The most commonly used are the normal distribution, Student *t*-distribution[2], and the Generalized Error Distribution (GED)[3].

Since it may be expected that excess kurtosis and skewness displayed by the residuals of conditional heteroscedasticity models will be reduced when a more appropriate distribution is used, we consider three distributions in this study: the Normal, the Student-t and the Generalized Error Distribution (GED).

[1] GARCH models can also be estimated by the Quasi Maximum Likelihood (QML) method introduced by (Bollerslev & Wooldridge, 1992) and by the Generalized Method of Moments (GMM) suggested and implemented by (Glosten, Jagannathan, & Runkle, 1993).

[2] Suggested by Bollerslev (1987); Baillie and Bollerslev (1989) and Beine, Laurent, and Lecourt (2000).

[3] Suggested by Nelson (1991) and Kaiser (1996).

Gaussian

The normal distribution is the most widely used when estimating GARCH models. The log-likelihood function for the standard normal distribution for the stochastic process of innovations given by (1) is represented as:

$$L_{normal} = -\tfrac{1}{2}\sum_{t=1}^{T}[\ln(2\pi) + \ln(\sigma_t^2) + z_t] \qquad (7)$$

where T is the number of observations.

Student-t

For a Student-t distribution, the log-likelihood is:

$$L_{student-t} = \ln\left[\Gamma\left(\frac{v+1}{2}\right)\right] - \ln\left[\Gamma\frac{v}{2}\right] - \frac{1}{2}\ln[\pi(v-2)] - \frac{1}{2}\sum_{t=1}^{T}\left[\ln\sigma_t^2 + (1+v)\ln\left(1+\frac{z_t^2}{v-2}\right)\right] \qquad (8)$$

where   is the degrees of freedom, $2 < \nu = \infty$ and $\Gamma(\cdot)$ is the gamma function.

Generalized Error Distribution (GED)

Skewness and kurtosis are important in financial applications in many aspects such as (in asset pricing models, portfolio selection, option pricing theory, Value-at-Risk and others). Therefore, a distribution that can model these two moments is appropriate, the GED log-likelihood function of a normalized random error is:

$$L_{GED} = \sum_{t=1}^{T}\left[\ln(v/\ \lambda_v) - 0.5\left|\frac{z_t}{\lambda_v}\right|^v - (1+v^{-1})\ln(2) - \ln\Gamma(1/v) - 0.5\ln(\sigma_t^2)\right] \qquad (9)$$

Where $\lambda_v = \sqrt{\frac{\Gamma(1/v\,2^{-2/v})}{\Gamma(3/v)}}$

and v is a positive parameter governing the thickness of the tails of the distribution. Note that for v=2, constant λ=1, the GED is equal to the standard normal distribution. For more details about the generalized error distribution, see Hamilton (1994).

DATA AND METHODOLOGY

Data

All data are the daily data obtained from DataStream. In the database, the daily return Rt consisted of daily stock closing price $P_{t,}$ which is measured in local currency[4] . Our measurements include Singapore's Strait Times Index (STI) and Malaysia's Kuala Lumpur Composite Index (KLCI).

[4] The stock returns were measured in local currency just as (K. H. Bae & Karolyi, 1994) and (K.-H. Bae & Andrew Karolyi, 1995) did in their studies. On the other hand, the stock returns in (Ng, 2000) is denominated in US dollars. Note that when market returns are denominated in US dollars, international investors are assumed to be unhedged against foreign exchange risk. However, (Dumas & Solnik, 1995) and (De Santis & Gerard, 1998) insist the importance of currency risk on stock markets. Thus, we assume that the investors are hedged against it.

The sample consisted of 3,652 daily observations on stock returns of the KLCI and the STI indices. It covers a fourteen-year period, beginning from 2 January 1991 and ending on 31 December 2004[5] . For illustrative purposes, Figure 1 compares the two indices' daily closing values taken across the sample period. Furthermore, Figure 2 looks at the behavior of the KLCI and STI returns, respectively, over the sample period. The data of stock price exhibit large fluctuations during the whole period. The indices prices are transformed into their returns so that we obtain stationary series. The transformation is as given below;

$$Rt = ln[(Pt) / ln(P_{t-1})] \qquad (10)$$

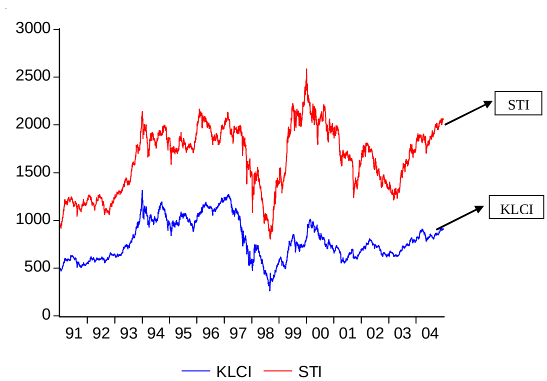

Fig.1: KLCI and STI Daily Closing Prices 2 January 1991- 31 December 2004

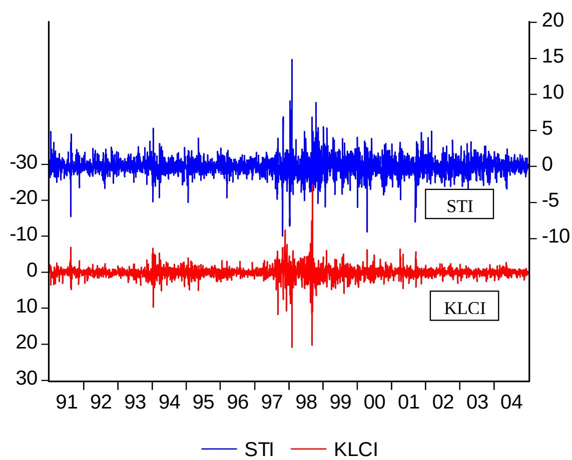

Fig.2: KLCI and STI Returns 2 January 1991- 31 December 2004

[5] All the data were supplied by Datastream.

The descriptive statistics of both indices in Table (1) over the sample period highlights the following:

- Mean returns for the STI Index is slightly larger than the KLCI, whereas, the non-conditional variance for the KLCI Index is larger than the STI. Furthermore, there is evidence of volatility clustering (See figure 2) and that large or small asset price changes tend to be followed by other large or small price changes of either sign (positive or negative). This implies that stock return volatility changes over time. Furthermore, the figures indicate a sharp increase in volatility starting from the year 1997.
- The returns for both indices are positively skewed. The null hypothesis for skewness coefficients that conforms to a normal distribution's value of zero has been rejected at the 5 percent significance level.
- The returns for both indices also display excess kurtosis. The null hypothesis for kurtosis coefficients that conform to the normal value of three is rejected for both indices.
- The high values of Jarque-Bera test for normality decisively rejects the hypothesis of a normal distribution.
- Moreover, Engle (1982) LM test indicates the presence of ARCH processes in the conditional variance. Both indices show signs of heteroskedasticity in the sample, indicating the legitimacy of using ARCH/GARCH type models.

Table 1: Summary Statistics for daily returns 1 January 1991-31 December 2004

	Sample	Mean	St. Dev.	Skewness	Ex-Kurtosis	Q(20)	$Q^2(20)$	J.Bera	ARCH(2)
KLCI	3652	0.0163	1.5731	0.5156	40.7437	105.69	1826.321	25255	580.8460
STI	3652	0.0216	1.2908	0.2884	11.2086	101.39	952.0316	19150	101.1180

J-Bera is the Jarque and Bera (1987) test for normality, ARCH(2) refers to Engle (1982) LM test for presence of ARCH at lag 2

The statistical results for both indices appear to have very similar characteristics. They both display positive skewness, were found to be deviating from normality, and display a degree of serial correlation. These stylized results are consistent with previous empirical work on the Asian-Pacific markets[6]and similar to a number of previous empirical works on matured markets[7] .

Finally, if we look at the sample, given the fact that the return series exhibited some excess kurtosis, it can also be predicted that a fatter-tailed distribution such as the student-

[6] See, (S. J. Kim, 2003), Ng. A. (2000).

[7] (Fama, 1976) showed that the distribution of both daily and monthly returns for the Dow Jones departs from normality, and are skewed, leptokurtic, and volatility clustered. Furthermore, (D. Kim & Kon, 1994) found the same for the S&P 500

t, or maybe a GED, should generate better results than just simply a normal distribution or a more complex asymmetric student-*t*.

EMPIRICAL RESULTS

Estimation and diagnostic

The quasi maximum likelihood approach is used to estimate the three models in equations (3), (4) and (6), with the three underlying error distributions. We consider ARCH (q) errors for $q = 1, 2, 3$, and 4, for the purpose of comparison. The lag length was selected according to the commonly used lag length selection criteria AIC and BIC. Low-order lag lengths were found to be sufficient to model the variance dynamics over very long sample periods[8] .

This section presents the estimation results and the validity, post estimation tests, of the estimated model. Tables 2, 3 and 4 present the estimation results for the parameters

Table 2: Estimation Statistics-Distributions Comparison AR(1)-GARCH Model

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

	Malaysia			Singapore		
	Normal	Student-t	GED	Normal	Student-t	GED
φ_0	0.04563	0.018806	0.005157	0.03876	0.019358	0..015127
	(0.0156)	(0.013)	(0.0152)	(0.0159)	(0.0147)	(0.0140)
φ_1	0.176251	0.150871	0.112706	0.135218	0.127392	0.090214
	(0.0179)	(0.0170)	(0.0237)	(0.0181)	(0.0173)	(0.0195)
α_0	0.021963	0.02101	0.036622	0.037439	0.030273	0.051645
	(0.00407)	(0.0049)	(0.00767)	(0.00657)	(0.00713)	(0.0104)
α_1	0.10164	0.078454	0.132121	0.127884	0.084825	0.0137722
	(0.00972)	(0.0107)	(0.0169)	(0.0125)	(0.0121)	(0.0175)
β_1	0.088967	0.842576	0.85267	0.854878	0.842505	0.833316
	(0.00967)	(0.0194)	(0.0165)	(0.0128)	(0.0209)	(0.0191)
v		4.241789	1.096725			1.24425
		(0.3307)	(0.0361)			(0.0396)

Asymptotic heteroskedasticity-consistent standard errors are given in parentheses.

[8] (French, Schwert, & Stambaugh, 1987) analyzed daily S&P stock index data for 1928-1984 for a total of 15,369 observations and required only four parameters in the conditional variance equation (including the constant).

Table 3: Estimation Statistics-Distribution Comparison AR(1)-EGARCH Model

$$R_t = \varphi_0 + \sum_{i=1}^{q} \phi_i R_{t-i} + \varepsilon_t$$

$$\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i g(z_{t-i}) + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-j}^2)$$

	Malaysia			Singapore		
	Normal	Student-t	GED	Normal	Student-t	GED
φ_0	0.039446	0.007716	0.001398	0.009178	-0.00026	0.000639
	(0.0155)	(0.0137)	(0.0124)	(0.0160)	(0.0429)	(0.0139)
φ_1	0.167528	0.148854	0.119731	0.136472	0.128629	0.092575
	(0.0171	(0.0165)	(0.233)	(0.0174)	(0.0176)	(0.0148)
α_0	0.011857	-0.04131	0.01857	0.014674	-0.03072	0.016111
	(0.00220	(0.00707)	(0.00425)	(0.00285)	(0.0113)	(0.00403)
α_1	0.160060	0.176339	0.198965	0.198497	0.188372	0.211591
	(0.0126)	(0.0197)	(0.0232)	(0.0163)	(0.0245)	(0.0233)
β_1	0.990215	0.974307	0.982273	0.978938	0.966408	0.973411
	(0.00188)	(0.00574)	(0.00421)	(0.00367)	(0.00751)	(0.00597)
g	-0.28794	-0.28878	-0.27707	-0.31462	-0.28956	-0.28998
	0.0453	(0.0523)	(0.0544)	(0.0448)	(0.0570)	(0.00543)
v		4.181703	1.1.7278		5.798267	1.261971
		(0.3161))	(0.0365)		(1.6173)	(0.0396)

Asymptotic heteroskedasticity-consistent standard errors are given in parentheses.

Table 4: Estimation Statistics-Distributions Comparison AR(1)-GJR Model

$$R_t = \varphi_0 + \sum_{i=1}^{q} \phi_i R_{t-i} + \varepsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} (\alpha_i \varepsilon_{t-1}^2 + w_i S_{t-i}^- \varepsilon_{t-1}^2) + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

	Malaysia			Singapore		
	Normal	Student-t	GED	Normal	Student-t	GED
φ_0	0.19303	0.003455	5.653E-6	0.013754	0.004819	0.00199
	(0.0160)	(0.0142)	(0.00165)	(0.163)	(0.0151)	(0.0140)
φ_1	0.18286	0.1566	0.116387	0.141688	0.12999	0.095103
	(0.0178)	(0.0167)	(0.0234)	(0.0178)	(0.0171)	(0.0181)
α_0	0.021062	0.022501	0.036813	0.036575	0.03107	0.050256
	(0.00374)	(0.00507)	(0.00751)	(0.00628)	(0.00703)	(0.00982)
α_1	0.13262	0.116718	0.185545	0.166968	0.120471	0.186991
	(0.0134)	(0.0161)	(0.0255)	(0.0169)	(0.0177)	(0.0242)
β_1	0.00922)	(0.0199)	(0.0163)	(0.0121)	(0.0202)	(0.0179)
ϖ_1	-0.8089	-0.07259	-0.11386	-0.10247	-0.07149	-0.11188
	(0.0125)	(0.0142)	(0.0232)	(0.0157)	(0.0156)	(0.0223)
v		4.329078	1.111723		5.829943	1.265056
		(0.3285)	(0.0366)		(0.5414)	(0.0404)

Asymptotic heteroskedasticity-consistent standard errors are given in parentheses.

for the GARCH, EGARCH and GJR-GARCH models respectively, while Tables 5-7 reports some useful in-sample statistics. Some comments can be made on these results:

- The use of asymmetric GARCH models seems justified. All asymmetric coefficients are significant at standard levels. Moreover, the Akaike information criteria (AIC) and the log-likelihood values highlight the fact that EGARCH or GJR models better estimate the series than the traditional GARCH.
- As is typical of GARCH model estimates for financial asset returns data, the sum of the coefficients on the lagged squared error ($\alpha 1$) and the lagged conditional variance (β_1) are close to unity (0.99 and 0.98) with the normal and GED error term for KLCI and STI respectively. This implies that shocks to the conditional variance will be highly persistent, indicating that large changes and small changes tend to be followed by small changes, this mean volatility clustering is observed in both KLCI and STI financial returns series.
- Regarding the densities (Tables 5-7), the symmetric distribution with fatter tails (Student-t and and GED distributions) clearly outperform the Gaussian. Indeed, the log-likelihood function strongly increases when using fatter tailed distribution. Using the Student-t leads to BIC criteria of 3.09 and 3.04 with the Gaussian versus 2.98 and 2.97 with the non-normal densities, for the KLCI and the STI respectively using AR(1)-GARCH, and similar results for both AR(1)-EGARCH and AR(1)-GJR .
- All the models seem to do a good job in describing the dynamics of the first two moments of the series as shown by the Box-Pierce statistics for the squared standardized residuals with lag 20 which are all non-significant at 5% level.
- LM test for presence of ARCH effects at lag 2, indicate that the conditional heteroskedasity that existed when the test was performed on the pure return series (see Table 1) are removed for GARCH but remains for EGARCH and GJR using

Table 5: Diagnostics statistics -Distributions Comparison AR(1)-GARCH Model

	Malaysia			Singapore		
	Normal	Student-t	GED	Normal	Student-t	GED
$Q^2(20)$	17.380	14.343	11.673	13.250	13.044	12.801
	(0.628)	(0.813)	(0.927)	(0.866)	(0.875)	(0.886)
ARCH(2)	2.924836	0.549214	0.032969	0.815931	1.128814	1.046473
	(0.053799)	(0.577451)	(0.967569)	(0.442308)	(0.323530)	(0.351280)
AIC	3.088865	2.969216	2.976918	3.044102	2.966946	2.968075
BIC	3.095661	2.977711	2.985413	3.050898	2.975440	2.976570
Log-Like	-5585	-5389	-5393	-5523	-5385	-5400

$Q^2(20)$ are the Box-Pierce statistic at lag 20 of the squared standardized residuals. P-values of are given in parentheses. AIC,BIC and Log-Like are the Akaike Information criterion, Swartz information criterion and Log-Likelihood value respectively.

Table 6: Diagnostics statistics -Distributions Comparison AR(1)-EGARCH Model

	Malaysia			Singapore		
	Normal	Student-t	GED	Normal	Student-t	GED
$Q^2(20)$	30.803	11.960	12.845	12.311	12.666	12.443
	(0.058)	(0.917)	(0.884)	(0.905)	(0.891)	(0.900)
ARCH(2)	10.01139	1.960507	2.301999	2.480221	1.647911	2.046106
	(0.000046)	(0.140935)	(0.100204)	(0.038866)	(0.192595)	(0.129386)
AIC	3.082735	2.961012	2.969516	3.035179	2.960053	2.960535
BIC	3.091229	2.971205	2.979709	3.043674	2.970246	2.970728
Log-Like	-5575	-5375	-5378	-5503	-5369	-5382

Table 7: Diagnostics statistics -Distributions Comparison AR(1)-GJR Model

	Malaysia			Singapore		
	Normal	Student-t	GED	Normal	Student-t	GED
$Q^2(20)$	20.424	13.503	11.944	12.496	18.124	12.550
	(0.432)	(0.855)	(0.918)	(0.898)	(0.579)	(0.896)
ARCH(2)	4.5840	1.1092	0.1454	0.814720	0.706989	1.460312
	(0.0103)	(0.3299)	(0.8647)	(0.442844)	(0.493194)	(0.232300)
AIC	3.076357	2.962434	2.967911	3.030473	2.958600	2.967269
BIC	3.086506	2.979803	2.970422	3.040646	2.970493	2.971731
Log-Like	-5561	-5370	-5375	-5497	-5369	-5384

$Q^2(20)$ are the Box-Pierce statistic at lag 20 of the squared standardized residuals. P-values of are given in parentheses. AIC,BIC and Log-Like are the Akaike Information criterion, Swartz information criterion and Log-Likelihood value respectively.

the Gaussian distribution. EGARCH and GJR models with student-t and GED distributions shows that the conditional heteroskedasity are successfully removed which are all non-significant at 5% level. From the previous, GARCH model performs better with Gaussian distribution. However, EGARCH and GJR models give better results with fatter tailed distributions.

- Similar to the results found in various markets, the leverage effect term $w_1<0$ in the GJR and g < 0 in the EGARCH are statistically significant at levels (p-value equal 0.01 and 0.05 respectively) with negative sign, as expected. The negative shocks imply a higher next period conditional variance than positive shocks of the same sign, indicating that the existence of leverage effect is observed in returns of the KLCI and STI stock market index.
- However, the comparison between models with each density (normal versus non-normal) shows that, according to the different measures used for modeling the volatility, the GJR-GARCH model with student-t provides the best in-sample estimation for KLCI having slight difference with EGARCH and clearly outperforms the symmetric models. No clear results can be obtained for the STI, where EGARCH and GJR with student-t provides a very close result (same Log-Like -5369). Figures 3 and 4 draws the behavior of the conditional variance for both models .

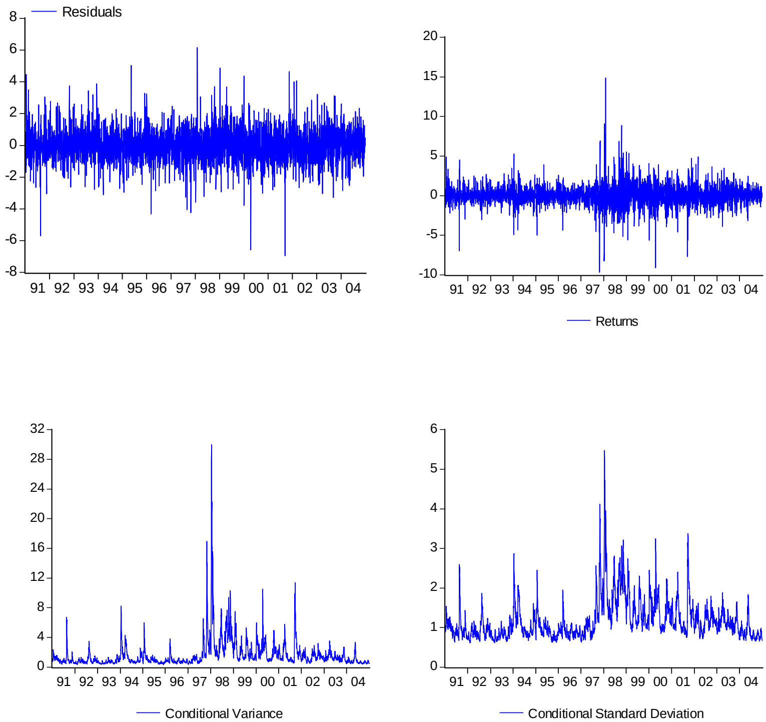

Fig.3:The STI Returns, Residuals and Conditional Variance AR(1)-GJR Model

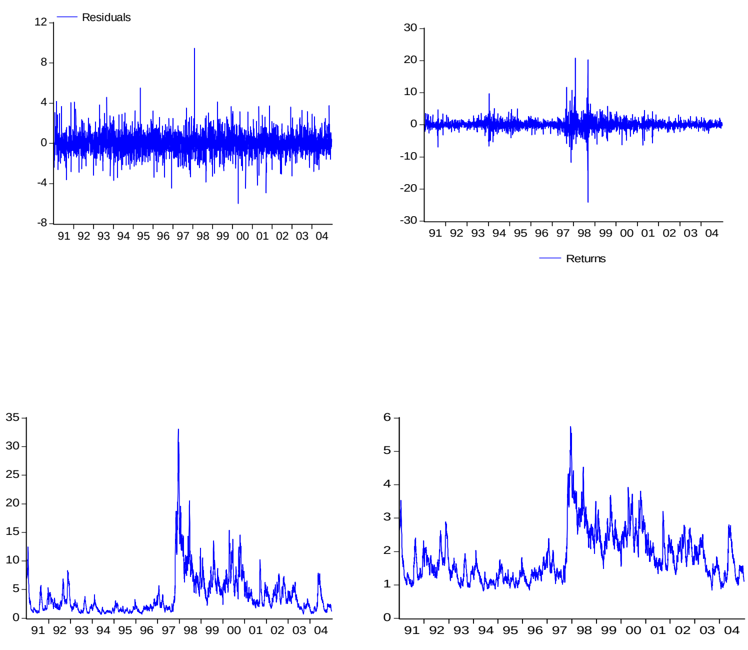

Fig.4: The KLCI Returns, Residuals and Conditional Variance AR(1)-GJR Model

Forecasting Evaluation

The one-step-ahead forecast of the conditional variance is easy to obtain. First we rewrite equation (3) and update h_t by one period,

$$h_{t+1} = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 h_t \tag{11}$$

Since ε_t^2 and h_t are known in period t, the one-step-ahead forecast is simply $\alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 h_t$. It is only somewhat more difficult to obtain the j-step-ahead forecasts. To begin, use the fact that $\varepsilon_t^2 = v_t^2 h_t$ so that $\varepsilon_{t+j}^2 = v_{t+j}^2 h_{t+j}$. If we update by j periods and take the conditional expectation of each side, it should be clear that $E_t \varepsilon_{t+j}^2 = E_t (v_{t+j}^2 h_{t+j})$. Since v_{t+j} is independent of h_{t+j} and $E_t v_{t+j}^2 = 1$, it follows that

$$E_t \varepsilon_{t+1}^2 = h_{t+1} \tag{12}$$

We can use (12) to obtain the forecasts of the conditional variance of the GARCH (1,1) process. Updating (11) by j periods, $h_{t+j} = \alpha_0 + \alpha_1 \varepsilon_{t+j-1}^2 + \beta_1 h_{t+j-1}$ and taking the conditional expectation,

$E_t h_{t+j} = \alpha_0 + \alpha_1 E_t \varepsilon_{t+j-1}^2 + \beta_1 E_t h_{t+j-1}$.If we combine this relationship with (12), it is easy to verify that $E_t h_{t+j} = \alpha_0 + (\alpha_1 + \beta_1) E_t h_{t+j-1}$ (13)

Given h_{t+1}, we can forecast all subsequent values of the conditional variance as:

$E_t h_{t+j} = \alpha_0 [1 + (\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)^2 + ... + (\alpha_1 + \beta_1)^{j-1}] + (\alpha_1 + \beta_1)^j h_t$.

If $\alpha_1 + \beta_1 < 1$, the conditional forecasts of h_{t-j} will converge to the long-run value $Eh_t = \alpha_0 / (1 - \alpha_1 - \beta_1)$.

Similarly, we can forecast the conditional variance of the ARCH (q) process using the following equation:

$$h_{t+1} = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + ... + \alpha_q \varepsilon_{t-q}^2 \tag{14}$$

Updating (14) by one period, $h_{t+1} = \alpha_0 + \alpha_1 \varepsilon_t^2 + ... + \alpha_q \varepsilon_{t-q+1}^2$

As mentioned above, at period t, we have all of the information necessary to calculate the value of h_{t+1} for any GARCH process. Now, if we update (14) by two periods and take the conditional expectation, we have $E_t h_{t+2} = \alpha_0 + \alpha_1 E_t \varepsilon_{t+1}^2 + ... + \alpha_q \varepsilon_{t-q+2}^2$

Since $E_t \varepsilon_{t+1}^2 = h_{t+1}$, it follows that $E_t h_{t+2} = \alpha_0 + \alpha_1 h_{t+1} + ... + \alpha_q \varepsilon_{t-q+2}^2$

It is clear from the preceding that it is possible to obtain the j-step-ahead forecasts of the conditional variance recursively. As the value of $j \to \infty$, the forecasts of h_{t+j} should converge to the unconditional mean $E\varepsilon_{t+1}^2 = \alpha_0 / (1 - \alpha_1 - \alpha_2 - \ldots - \alpha_q)$.

It should be clear that a necessary condition for convergence is for the roots of the inverse characteristics equation $1 - a_1 L - \ldots - a_q L^q$ to lie outside the unit circle. This is a necessary condition for the long-run mean to have the representation $a_0 / (1 - \Sigma a_i)$. To ensure that the variance is always positive, it is necessary that $a_0 > 0$ and $\alpha_i \geq 0$ for $i \geq 1$.

Since the conditional variance has been estimated, the obvious question is how good are the models for forecasting future conditional variance? Typically, there are several plausible models that we can select to use for our forecast. It is fallacious to conclude that the one with the best fit is the one that gives the best forecast. To assess the performance of the GARCH models candidates in forecasting the conditional variance, we compute 5 statistical measures of fit:

- Mean Squared Errors (MSE)
- Mean Absolute Error (MAE)
- Mean Absolute Percentage Error (MAPE)
- Theil Inequality Coefficient (TIC)
- Amomiya Prediction Criterion (APC)

The MSE is represented as:

$$\frac{1}{h+1}\sum_{t=s}^{s+h}(\hat{\sigma}_t^2 - \sigma_t^2)^2$$

where h is the number of steps ahead, (in this paper h is equal to 1, representing one step ahead), s the sample size, $\hat{\sigma}^2$ the forecasted variance and σ^2 is the conditional variance estimated from equations (3), (4) and (6)..

The MAE is:

$$\frac{1}{h+1}\sum_{t=s}^{s+h}|\hat{\sigma}_t^2 - \sigma_t^2|$$

The MAPE is represented as:

$$\frac{1}{h+1}\sum_{t=s}^{s+h}\left|\left(\hat{\sigma}_t^2 - \sigma_t^2\right)/\sigma_t^2\right|$$

The TIC is represented as:

$$TIC = \frac{\sqrt{MSE}}{\sqrt{1/h+1\sum_{t=s}^{s+h}\sigma^2} + \sqrt{1/h+1\sum_{t=s}^{s+h}\hat{\sigma}^2}}$$

Theil inequality coefficient is a scale invariant measure that always lies between zero and one, where zero indicates a perfect fit.

The APC, Amemiya prediction criterion is defined as:

$$APC = \left(\frac{s+k}{s-k} \right) \frac{1}{s} SSE$$

The results of forecasting daily volatility with GARCH models together with various distributions and five evaluation criteria are given in Table 8. All results are presented for each distribution and for each stock market. Table 8 reports detailed results about the forecast accuracy analysis based on classical evaluation criteria. Each section of the Table refers to a specific stock index, whereas the whole set of evaluation criteria is applied to each GARCH models, whose specification is always of order (1,1), and calculated for each of the three distributions. In this study, the length of the out-of-sample period was chosen to be 360 days.

Table 8: Forecast Performance out-of- Sample

	KLCI			STI		
			Normal			
	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR
MSE	2.0211	0.5938	0.5625	0.5910583	0.1779	0.2945
MAE	0.2933	0.2113	0.2180	0.2445678	0.1906	0.2116
MAPE	12.50454	9.8018	9.9522	13.413375	11.3728	12.0916
TIC	0.1952	0.1581	0.1485	0.2647	0.1811	0.2088
APC	2.38595	0.6632	0.6394	0.5892	0.1788	0.1659
R^2	0.94893	0.9646	0.96869	0.89148	0.93582	0.92575
			Student-t			
	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR
MSE	3.9085	1.4998	1.3143	0.5305	0.2066152	0.2821
MAE	0.4081	0.3110	0.3118	0.2317	0.203192	0.2167
MAPE	16.4260	13.9788	13.5583	12.7713	12.177492	12.4356
TIC	0.2391	0.2345	0.2011	0.2608	0.2002378	0.2108
APC	4.5453	1.6143	1.4630	0.5239	0.2072	0.2833
R^2	0.92587	0.92665	0.9552	0.89304	0.92042	0.92251
			GED			
	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR
MSE	6.0594	1.3041	1.2848	0.5305	0.1877	0.2805
MAE	0.5086	0.2894	0.3041	0.2317	0.1944	0.2123
MAPE	21.2935	13.6436	13.8632	12.7713	11.7055	12.2234
TIC	0.2931	0.2307	0.2067	0.2608	0.1911	0.2097
APC	6.9381	1.4029	1.42064	0.5287	0.1884	0.2818
R2	0.89186	0.92893	0.94298	0.89262	0.92743	0.92387

MSE is Mean Squared Error, MAE is the Mean Absolute Error, MAPE is the Mean Absolute Percentage Error, TIC is the Theil Inequality Coefficient, APC is Amemiya Prediction Criterion and R^2 is the Amemiya Adjusted R^2.

Table 9 gives the rank of the GARCH models (when evaluated against each other) with the three different distributions for the error term. From Tables 8 and 9, some interesting observations and conclusions emerge. A first major conclusion from all the Tables is that there is no single model that completely dominates the other models for both series. Secondly, forecasting with normal distribution does not yield a significant reduction of the forecast error relative to the GED and student-t distribution. Thus, the failure of predictor $\hat{\sigma}_t^2$ is justified due to the fact that the GARCH model's residuals follow a (possibly) heavy-tailed distribution. Third, it seems that asymmetric models (EGARCH and GJR) with a fatter tailed distribution tend to produce better forecast. It is apparent that the simple predictor $\hat{\sigma}_t^2$ seems to actually have some predictive ability, when a heavy-tailed is assumed for the GARCH residuals.

A general finding is that the asymmetric models EGARCH and GJR-GARCH model with heavy-tailed distribution are the best performers while the GARCH model is the worst. A possible explanation is that modeling asymmetries contributes to the reduction of the magnitude of the bias. For KLCI, there is an indication that GJR-GARCH model

Table 9: Ranking Forecast Performance

	KLCI			STI		
	Normal					
	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR
MSE	3	2	1	3	1	2
MAE	3	1	2	3	1	2
MAPE	3	1	2	3	1	2
TIC	3	2	1	3	1	2
APC	3	2	1	3	2	1
Total	**15**	**8**	**7**	**15**	**6**	**9**
	Student-t					
	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR
MSE	3	2	1	3	1	2
MAE	3	1	2	3	1	2
MAPE	3	2	1	3	1	2
TIC	3	2	1	3	1	2
APC	3	2	1	3	1	2
Total	**15**	**9**	**6**	**15**	**5**	**10**
	GED					
	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR
MSE	3	2	1	3	1	2
MAE	3	1	2	3	1	2
MAPE	3	1	2	3	1	2
TIC	3	2	1	3	1	2
APC	3	1	2	3	1	2
Total	**15**	**7**	**8**	**15**	**5**	**10**

with student-t distribution performs better than EGARCH. However, EGARCH with a heavy-tailed distribution performed better than GJR model for STI.

Moreover, Table 8 shows that, the R2 is higher when using asymmetric GARCH. For instance, when using a Student-t distribution, it ranges from 0.927 to 0.946 with the asymmetric GARCH versus 0.925 with the symmetric GARCH for the KLCI and it goes from 0.920 to 0.923 versus 0.893 with the symmetric GARCH for the STI. A description of the fitted and forecasted variance of EGARCH and GJR models is shown in Figures 5 and 6.

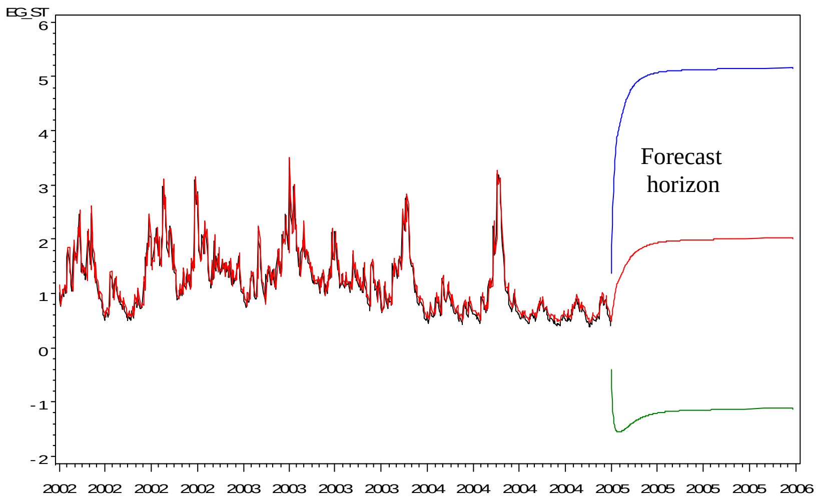

Fig 5: EGARCH, the fitted and forecasted variance, estimated through 2005 for STI

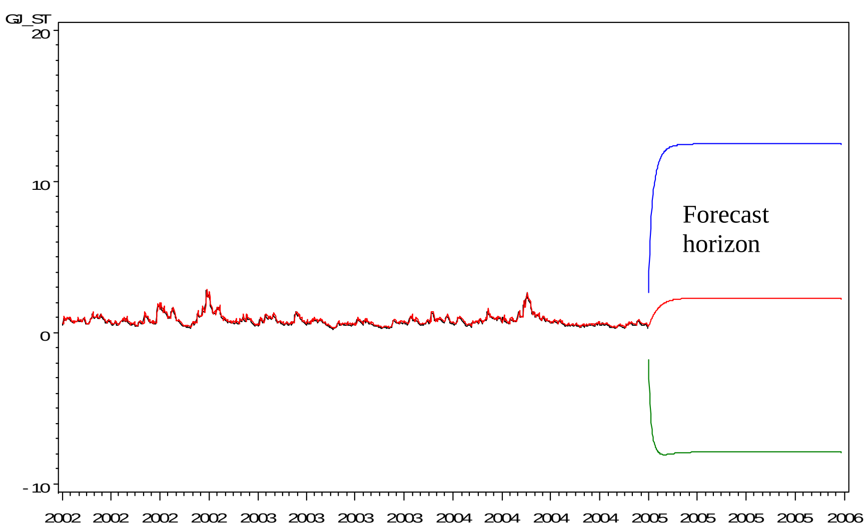

Fig 6: GJR, the fitted and forecasted variance, estimated through 2005 for KLCI

CONCLUSION

The volatility of stock prices has received great attention from both academies and practitioners over the last two decades because it can be used as a measure of risk in financial markets. Recent portfolio selection, asset pricing, value at risk, option pricing and hedging strategies, highlight the importance of modeling and forecasting the conditional volatility of returns.

This paper contributes to the literature of volatility modeling in two ways. First, data set was used from an emerging market. Secondly, the alternative ARCH-type models (symmetric and asymmetric GARCH Models) were estimated. The comparison was focused on two different aspects: the difference between symmetric and asymmetric GARCH (i.e., GARCH versus EGARCH and GJR-GARCH) and the difference between normal tailed symmetric, fat-tailed symmetric distributions (i.e. Normal versus Student-t and Generalized Error Distribution) for estimating the KLCI and STI stock market index returns volatility.

The in-sample statistical results indicate that the estimated parameters of the AR(1)-GJR model, the coefficients of ARCH(α_1) and GARCH(β_1) in the conditional variance equation of the AR(1)-GJR in both markets are highly significant with p-value equal to 0.016 and 0.019 for KLCI , 0.017 and 0.020 for STI.

As expected with the results found in various markets, the leverage effect term (w_1) in both KLCI and STI markets, the AR(1)-GJR Model is statistically significant at levels (p-value equal 0.014 and 0.015 respectively) with a negative sign, which indicate that negative shocks imply a higher next period conditional variance than positive shocks of the same sign, indicating that the existence of leverage effect is observed in returns of the KLCI and STI stock market index.

However, the comparison between models with each density (normal versus non-normal) shows that, according to the different measures used for the performance of volatility forecast, the GJR-GARCH model provides the best out-sample estimation for KLCI and EGARCH model provides the best out-sample estimation for STI, and clearly the asymmetric models outperform symmetric models. Our results show that noticeable improvements can be made when using a GARCH model in the conditional variance (and, among the tested models, EGARCH and GJR seem to outperform GARCH). Moreover, non-normal distributions provide better in-sample results than the Gaussian distribution. However, out-of-sample results show less evidence of superior forecasting ability.

In general, from the results, we can argue that the asymmetric models (GJR and EGARCH model) coupled with a Student-t distribution for the innovations, performs very well with the dataset investigated. The models seem to capture the dynamics of the first and second moments of the KLCI and STI stock market index returns series.

Finally, future research could be directedb at forecasting the volatility of the KLCI and STI financial time series. First, "true volatility" could be better estimated by selecting shorter time intervals (for instance, intra-day trading). Second, introducing long run persistence shocks in the volatility with fractionally integrated models (FIGARCH, FIEGARCH, FIAPARCH) would certainly allow better insights into the dynamics of the series.

REFERENCES

BAE, K.-H. and G. ANDREW KAROLYI. 1995. Good news, band news and international spilovers of stock return volatility between Japan and the U.S. *Pacific-Basin Finance Journal* **3(1):** 144-144.

BAE, K. H. and G. A. KAROLYI. 1994. Good news, bad news and international spillovers of stock return volatility between Japan and the US.

BOLLERSLEV, T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31(3):** 307-327.

BOLLERSLEV, T. and J. WOOLDRIDGE. 1992. Quasi maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Reviews* **5.**

CHOUDHRY, T. 2005. Time-varying beta and the Asian financial crisis: Evidence from Malaysian and Taiwanese firms. *Pacific-Basin Finance Journal* **13(1)**: 93-118.

DE SANTIS, G.. and B. GERARD. 1998. How big is the premium for currency risk? *Journal of Financial Economics* **49(3):** 375-412.

DUMAS, B. and B. SOLNIK. 1995. The world price of foreign exchange rate risk. *Journal of Finance* **50(2):** 445-479.

ENGLE, R. F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50:** 987-1006.

FAMA, E. F. 1976. *Foundations of Finance: Portfolio Decisions and Securities Prices: Basic Books.*

FRENCH, K. R., G. W. SCHWERT and R. STAMBAUGH. 1987. Expected stock return and volatility. *Journal of Financial Economics* **19:** 3-29.

GLOSTEN, L., R. JAGANNATHAN and D. RUNKLE. 1993. Relationship between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* **48(1779):** 1801.

KIM, D. and S. J. KON. 1994. Alternative models for the conditional heteroscedasticity of stock returns. *The Journal of Business* **67(4):** 563-598.

KIM, S. J. 2003. The spillover effects of US and Japanese public information news in advanced Asia-Pacific stock markets. *Pacific-Basin Finance Journal* **11(5).**

KON, S. J. 1984. Models of stock returns--A comparison. *The Journal of Finance* **39(1):** 147-165.

LAMBERT, P. and S. LAURENT. 2001. Modelling financial time series using GARCH-type models with a skewed Student distribution for the innovations. Institut de Statistique, Louvain-la-Neuve Discussion Paper, 125.

Liu, S.M. and B.W. Brorsen. 1995. Maximum likelihood estimation of a GARCH-stable model. *Journal of Applied Econometrics* **10(3):** 273-285.

Mittnik, S. and M.S. Paolella. 2001. *Prediction of Financial Downside-Risk with Heavy-Tailed Conditional Distributions. Handbook of Heavy Tailed Distributions in Finance.*

Nelson, D.B. 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometric*a **59(2):** 347-370.

Ng, A. 2000. Volatility spillover effects from Japan and the US to the Pacific-Basin. *Journal of International Money and Finance* **19(2):** 207-233.

Pan, M. S., Y. A. Liu and H. J. Roth. 1999. Common stochastic trends and volatility in Asian-Pacific equity markets. *Global Finance Journal* **10(2):** 161-172.

Simkowitz, M. A. and W. L. Beedles. 1980. Asymmetric stable distributed security returns. *Journal of the American Statistical Association* **75(370):** 306-312.

Tsay, R.S. 2002. *Analysis of Financial Time Series.* New York: Wiley.

Wen Cheong, C., A. Hassan Shaari Mohd Nor and Z. Isa. 2007. Asymmetry and long-memory volatility: Some empirical evidence using GARCH. *Physica A: Statistical and Theoretical Physics* **373:** 651-664.

Zakoian, J.M. 1994. Threshold heteroskedasticity models. *Journal of Economic Dynamics and Control* **15:** 931-955.