Statistical Inadequacy of GARCH Models for Asian Stock Markets: Evidence and Implications

Kian-Ping Lim
Melvin J. Hinich
Venus Khim-Sen Liew

This study employs the Hinich portmanteau bicorrelation test (Hinich 1996; Hinich and Patterson 1995) as a diagnostic tool to determine the adequacy of Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models for eight Asian stock markets. The bicorrelation test results demonstrate that this type of model cannot provide an adequate characterisation for the underlying process of all the selected Asian stock markets. Further investigation using the windowed test procedure reveals that the violation of the covariance stationarity assumption as required by the GARCH process is due to the presence of transient epochs of dependencies in the data. The inadequacy of GARCH models has strong implications for the pricing of stock index options, portfolios selection, development of optimal hedging techniques and risk management.

JEL Classification: G120, C520
Keywords: GARCH, non-stationarity, data generating process, bicorrelation, Asian stock markets

1. Introduction

Time series models have been widely employed in the literature to explain the dynamics of financial time series. Since its introduction 22 years ago, applications of the Autoregressive Conditional Heteroscedasticity (ARCH) model introduced by Engle (1982) or its extension GARCH by Bollerslev (1986)

JOURNAL OF EMERGING MARKET FINANCE, 4:3 (2005)
Sage Publications New Delhi ■ Thousand Oaks ■ London
DOI: 10.1177/097265270500400303

in finance have become commonplace (for a survey, see Bollerslev et al. 1992; Engle 1995). This popularity is evidenced by the incorporation of GARCH estimation into major software packages (for a review, see, for example, Brooks 1997; Brooks et al. 2001; McCullough and Renfro 1999). Generally, this type of non-linear time series model is able to capture a special type of non-linearity in the data generating process, known as multiplicative non-linearity, or non-linear-in-variance, in which non-linearity affects the process through its variance (Hsieh 1989).

In the literature, the family of GARCH models has grown at a tremendous rate. Engle (1995), Hentschel (1995) and Pagan (1996), among others, provided an excellent account of the variations and extensions of GARCH models over the years. However, in most applications, GARCH (1, 1) model is found to suffice. According to Brooks (1996: 309), it is highly unlikely that a GARCH model of order greater than one in the autoregressive and moving average components would be required. This is because a GARCH (1, 1) model implies an infinitely long memory with respect to past innovations.

The simple GARCH (1, 1) model for $\{y(t)\}$ can be written as:

$$\begin{aligned} & y(t) = \varepsilon_t h_t^{1/2} \\ & \varepsilon_t \mid \psi_{t-1} \sim N(0, h_t) \\ & h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \end{aligned} \tag{1}$$

To ensure a well defined process in the context of GARCH (1, 1) model, the following customary constraints are applied to the parameters: $\alpha_1 \geq 0$, $\beta_1 \geq 0$ and $\alpha_1 + \beta_1 < 1$ suffices for covariance stationarity.[1] In equation (1), ε_t is conditional on the information set ψ_{t-1} and is normally distributed with zero mean and variance h_t. In the GARCH models, the variance is time varying and this provides an alternative and useful measurement of volatility.[2] For most financial data, one of the stylised features is that they do experience volatility clustering; that is, there are stable periods with only minor adjustments and unstable periods with large deviations for many consecutive days. GARCH models have been popular because they are able to capture this

[1] A time series x_t is covariance stationary if (*a*) the mean of x_t is constant over time; (*b*) the variance of x_t is also constant over time; and (*c*) the covariance between x_t and x_{t+h} depends only on their time difference or lag, represented by h.

[2] Earlier studies have used unconditional standard deviation or variance as the measure of volatility and the limitations of such measure are well documented. See, for example, Jansen 1989; Pagan and Ullah 1988.

clustering feature. Despite their popularity, the ARCH parameterisation of the conditional variance does not have any solid grounding in economic theory, but represents 'a convenient and parsimonious representation of the data' (Hall et al. 1989).

Though the GARCH models have been known for more than two decades, most of their applications are on the widely traded financial markets of developed industrialised countries. However, there is a growing trend of their applications in the Asian stock markets, for instance forecasting volatility of stock returns (Choo et al. 1999), modelling the volatility of stock index futures market (Tan 2001) and determining volatility spillover effects among major Asia-Pacific stock markets (Hooy and Tan 2002). While these Asian studies involved different applications of GARCH models, none of them conducted a thorough investigation to determine the adequacy of the GARCH models. This issue is of great importance to the field of finance in view of the wide application of GARCH models in understanding the relationship between risk[3] and expected returns, particularly in the areas of asset pricing, portfolio selection and risk management. Motivated by the above consideration, this study contributes to the current literature by addressing the fundamental issue of GARCH adequacy in characterising the behaviour of major Asian stock returns series.

In terms of methodology, various procedures have been utilised or proposed in the literature to test whether the GARCH formulation represents an adequate characterisation of the data. It has been suggested that the Brock-Dechert-Scheinkman test (hereafter denoted as the BDS test) developed in Brock et al. (1987, 1996) can be used as a general test of model misspecification. Since the BDS test has reasonable power against the GARCH models, it has been utilised as a diagnostic tool to determine the adequacy of GARCH models for the detected non-linear structure. In this case, the standardised residuals from the fitted GARCH model are subjected to the BDS test. It is important to note that the null hypothesis for the test has now become 'the fitted GARCH model sufficient to model the detected non-linear structure in the data' against an unspecified alternative that it is not. If the BDS test cannot reject the null using appropriate critical values derived from simulation, then the fitted GARCH model is assumed to be an adequate characterisation of the data. This procedure has been followed by various authors in the literature (see, for example, Abhyankar et al. 1995; Hsieh 1989, 1991;

[3] Volatility of returns is considered by researchers and investors to be a proxy for risk. In this aspect, the GARCH estimates of time-varying volatility have been widely utilised in the financial world.

Krager and Kugler 1993; McMillan and Speight 2001; Opong et al. 1999). The main reason the BDS test has been popular among researchers in the earlier years is that the computer programmes for implementing the test have been made available and distributed widely.[4]

However, the problem with using the BDS test in this manner arises because the asymptotic distribution could not approximate very well the BDS statistics applied to standardised residuals of ARCH, GARCH and EGARCH (Brock et al. 1991; Hsieh 1991). The Monte Carlo results from Brooks and Heravi (1999) provided further evidence against the sole use of the BDS test as a general misspecification test for GARCH models. Specifically, in the simulation exercise, the authors generated samples of data by an AR(1), MA(1), self-exciting threshold autoregressive (SETAR) and bilinear models respectively. A GARCH model is estimated for each set of the generated data. Subsequently, the BDS statistics are calculated on the standardised residuals of this fitted GARCH model. As a whole, the results demonstrate that the effect of misspecified GARCH filter is more pronounced when the data generating process is non-linear, indicating that the BDS test could not reliably discriminate among different types of non-linear structure (in this case, the bilinear, SETAR and a GARCH model). Hence, the authors warned that it might not be appropriate to use non-rejection by the BDS test on the standardised residuals of a GARCH model as evidence that the GARCH model 'fits' the data.

Some authors have suggested that the joint application of existing non-linearity tests can provide deeper insight into the nature of non-linear generating mechanism of a time series (see, for example, Ashley and Patterson 2001; Barnett et al. 1995, 1997; Barnett and Serletis 2000). Following this recommendation, Lim (2001) utilised the differing power of the BDS and Hinich bispectrum tests (Hinich 1982) in detecting GARCH as an alternative framework for determining the adequacy of the GARCH models in characterising the behaviour of the series under study. In this case, the low power of the Hinich bispectrum test relative to the BDS test for GARCH models suggests that the bispectrum test is useful as a marker for the GARCH models. However, this approach has not received much attention from other researchers, possibly due to the concern raised by Psaradakis (2001) that there is an increased risk of overstating the significance of non-linearity when multiple tests are carried out using the same data.

[4] For instance, the code written for DOS-based computers was first provided by W.D. Dechert in his web page. Later, LeBaron (1997) shared the source code in the C programming language and provided a brief description of the BDS algorithms. In a recent development, the BDS test has been incorporated in the statistical package of *EViews*, starting from version 4.0.

The Hinich portmanteau bicorrelation test (Hinich 1996; Hinich and Patterson 1995), which was designed to detect episodes of transient serial dependencies within a data series, is an alternative technique to check the adequacy of GARCH models as valid data descriptions. From the Monte Carlo simulations given in Hinich and Patterson (1995), it was demonstrated that the bicorrelation test has better small sample properties and does not have the stiff data requirements of many of its competitors, such as the BDS test. Several recent studies (see, for example, Brooks and Hinich 1998; Brooks et al. 2000; Hinich and Patterson 1995; Lim et al. 2003) have utilised the bicorrelation test to examine the validity of specifying a GARCH error structure.

This study utilises the Hinich portmanteau bicorrelation test as a diagnostic tool to determine the adequacy of GARCH models in characterising the behaviour of eight Asian stock markets. The focus of this article is on the assumption of strict stationarity of GARCH models, which can be tested using the bicorrelation test. This line of inquiry is highly relevant since the issue of GARCH adequacy in these Asian stock markets has hardly received a mention in the literature.

This article provides in the following section a brief description of the data used in this study. Section 3 discusses the Hinich portmanteau bicorrelation test. Section 4 presents the empirical results as well as the analysis of the findings. Finally, implications and concluding remarks are given at the end of the article.

2. The Data

In this study, the data consist of daily closing prices for eight major Asian stock market indices: Bangkok S.E.T. (Thailand), Hang-Seng (Hong Kong), Jakarta SE Composite (Indonesia), Korea SE Composite (South Korea), Kuala Lumpur SE Composite (Malaysia), Nikkei 225 Stock Average (Japan), Philippines SE Composite (the Philippines) and Singapore Straits Times (Singapore). All these indices collected from *Datastream* are denominated in their respective local currency units. The prices covering the sample period from 2 January 1990–31 December 2003 are transformed into a series of continuously compounded percentage returns, using the relationship:

$$r_t = 100^* \ln(P_t / P_{t-1}) \qquad (2)$$

where P_t is the closing price of the stock on day t, and P_{t-1} the price on the previous trading day.

3. Hinich Portmanteau Bicorrelation Test

This section provides a brief description of the Hinich portmanteau bicorrelation test. A full theoretical derivation of the test statistics and a number of Monte Carlo simulations to assess their size and power are given in Hinich and Patterson (1995) and Hinich (1996).

Let the sequence $\{x(t)\}$ denote the sampled data process, where the time unit, t, is an integer. The test procedure employs non-overlapped data window, thus if n is the window length, then the *k-th* window is $\{x(t_k)$, $x(t_k + 1)$, ... $x(t_k + n-1)\}$. The next non-overlapped window is $\{x(t_{k+1})$, $x(t_{k+1} + 1)$, ... $x(t_{k+1} + n-1)\}$, where $t_{k+1} = t_k + n$. The null hypothesis for each window is that $x\{t\}$ are realisations of a stationary pure noise process[5] that has zero bicovariance. The alternative hypothesis is that the process in the window is random with some non-zero correlations $C_{xx}(r) = E[x(t) \times (t + r)]$ or non-zero bicorrelations $C_{xxx}(r, s) = E[x(t) \times (t + r) \times (t + s)]$ in the set $0 < r < s < L$, where L is the number of lags.

We can define $Z(t)$ as the standardised observations obtained as follows:

$$Z(t) = \frac{x(t) - m_x}{s_x} \tag{3}$$

for each $t = 1, 2, \ldots, n$ where m_x and s_x are the sample mean and sample standard deviation of the window. The sample correlation is:

$$C_{zz}(r) = (n - r)^{-\frac{1}{2}} \sum_{t=1}^{n-r} Z(t)Z(t + r) \tag{4}$$

The C statistic, which is developed for the detection of linear serial dependencies within a window, is defined as:

$$C = \sum_{r=1}^{L} [C_{zz}(r)]^2 \sim \chi^2_{(L)} \tag{5}$$

[5] A stationary time series is called pure noise or pure white noise if $x(n_1), \ldots n_N$. A white noise time series, by contrast, is one for which the autocovariance function is zero for all lags. Whiteness does not imply that $x(n)$ and $x(m)$ are independent for $m \neq n$ unless the series is Gaussian.

The (*r*, *s*) sample bicorrelation is:

$$C_{zzz}(r,s) = (n-s)^{-1}\sum_{t=1}^{n-s} Z(t)Z(t+r)Z(t+s) \qquad \text{for } 0 \le r \le s \tag{6}$$

The *H* statistic, which is developed for the detection of non-linear serial dependencies within a window, is defined as:

$$H = \sum_{s=2}^{L}\sum_{r=1}^{s-1} G^2(r,s) \sim \chi^2_{(L-1)(L/2)} \tag{7}$$

where $G(r,s) = (n-s)^{\frac{1}{2}} C_{zzz}(r,s)$.

In both the *C* and *H* statistics, the number of lags *L* is specified as $L = n^b$ with $0 < b < 0.5$, where *b* is a parameter under the choice of the user. Based on the results of the Monte Carlo simulations, Hinich and Patterson (1995) recommended the use of $b = 0.4$ in order to maximise the power of the test while ensuring a valid approximation to the asymptotic theory.

A window is significant if either the *C* or *H* statistic rejects the null of pure noise at the specified threshold level. This study uses a threshold of 0.01. In this case, the chance of obtaining a false rejection of the null is approximately one out of every 100 windows. With such a low-level threshold, it would minimise the chance of obtaining false rejections of the null hypothesis indicating the presence of dependencies where these do not exist.

It is possible to use the bicorrelation test to examine whether a GARCH formulation represents an adequate characterisation of the data. This is achieved by transforming the returns into a set of binary data, where

$$\{y(t)\}: \; y(t) = 1 \;\text{ if } Z(t) \ge 0$$

$$y(t) = -1 \;\text{ if } Z(t) < 0 \tag{8}$$

If $Z(t)$ are generated by a GARCH process whose innovations $\{\varepsilon_t\}$ are symmetrically distributed around a zero mean, then the binary set $\{y(t)\}$ will be a stationary pure noise series. The justification for the binary transformation is that it turns a GARCH into a pure noise.[6] Putting it differently, a GARCH process that has symmetric innovations produces independently distributed binary output. The binary transformed data have moments that are well behaved with respect to the asymptotic theory. Therefore, if the null of pure noise is rejected by the C or H statistics, then there are statistical structures present in the data that cannot be captured by GARCH models. The rejection may be due to serial dependence in the innovations but this violates a critical assumption for GARCH models.

4. Empirical Results

Table 1 provides summary statistics for all the Asian stock returns series under study. The mean values indicate that the daily returns for these Asian stock markets are on average very small. Hong Kong Hang-Seng recorded the highest returns of all with 0.04 per cent. In contrast, Bangkok S.E.T., Korea SE Composite and Nikkei 225 Stock Average performed poorly on average, with negative returns being reported. This poor average performance is associated with higher risk, especially for Bangkok S.E.T. and Korea SE Composite, as indicated by higher values of standard deviations for these two markets. Moving beyond the basic mean and standard deviation measurements to higher order moments, all of the returns series, with only two exceptional cases, exhibit some degree of positive or right skewness. On the other hand, the distributions for all the series are highly leptokurtic, in which the tails of their respective distributions taper down to zero more gradually than do the tails of a normal distribution. Not surprisingly, given the non-zero skewness levels and excess kurtosis, the Jarque-Bera (JB) test statistics clearly indicate that all the returns series significantly deviate from normality.

Table 2 presents the correlations (C) and bicorrelations (H) test statistics[7] for the binary transformed data set $\{y(t)\}$ covering the full sample period. The results show that the null of pure noise is strongly rejected by both the C and H statistics in all the eight Asian stock markets under investigation, and hence violate the assumption of covariance stationary of GARCH models. In other words, the GARCH models cannot provide an adequate

[6] Though ARCH/GARCH is a martingale difference process, and thus white noise, it is not pure noise.

[7] Both the C and H test statistics in this study are computed using the T23 programme written by the second author. The programme is available and any request should be directed to hinich@mail.la.utexas.edu.

Table 1
Summary Statistics

	BSET	*HKHS*	*JSE*	*KLSE*	*KSE*	*NIKKEI*	*PSE*	*SST*
Sample period	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003
No of observations	3,652	3,652	3,652	3,652	3,652	3,652	3,652	3,652
Mean	–0.003555	0.040777	0.015073	0.009447	–0.003155	–0.035415	0.007306	0.011345
Median	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Maximum	11.34953	17.24711	13.12768	20.81737	10.02377	12.43033	16.17760	14.86849
Minimum	–10.02803	–14.73471	–12.73214	–24.15339	–12.80470	–7.233984	–9.744158	–9.671880
Std deviation	1.804396	1.640428	1.520482	1.605225	1.967787	1.500535	1.591777	1.327327
Skewness	0.244347	–0.028993	0.339348	0.469388	–0.012413	0.214704	0.547349	0.215886
Kurtosis	7.648661	12.81171	14.17528	40.69354	6.743449	6.311444	11.85507	13.40801
JB normality test statistic	3,324.670	14,649.53	19,073.72	216,332.9	2,132.467	1,696.666	12,114.08	16,512.08
(*p*-value)	0.000000*	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

Notes: BSET—Bangkok S.E.T.; HKHS—Hong Kong Hang-Seng; JSE—Jakarta SE Composite; KLSE— Kuala Lumpur SE Composite; KSE—Korea SE Composite; NIKKEI—Nikkei 225 Stock Average; PSE— Philippines SE Composite; SST—Singapore Straits Times.
*denotes extremely small *p*-value.

Table 2

C and H Statistics for Whole Sample of Binary Trasformed Data

	BSET	*HKHS*	*JSE*	*KLSE*	*KSE*	*NIKKEI*	*PSE*	*SST*
Sample period	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003	2/01/1990–31/12/2003
No of observations	3,652	3,652	3,652	3,652	3,652	3,652	3,652	3,652
No of lags	27	27	27	27	27	27	27	27
p-value	0.0000*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
- C Statistic	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
- H Statistic								

Note: * denotes extremely small p-value.

characterisation for the underlying process of all the selected Asian stock markets since there are statistical structures present in the data that cannot be captured by this type of model.

The next step is to investigate the persistency of the underlying non-stationarity. One approach is to apply the bicorrelation test to smaller sub-samples, in which the whole sample period is divided into relatively shorter equal length windows of time. In this regard, there is no unique value for the window length. According to Brooks and Hinich (1998), the window length should be sufficiently long to provide adequate statistical power and yet short enough for the test to be able to pinpoint the arrival and disappearance of transient dependencies, which is the main purpose of using a windowed test procedure. In this study, we follow the choice of Brooks and Hinich (1998) in which the data are split into a set of non-overlapping windows of 35 observations in length, approximately seven trading weeks. In fact, it was found that the choice of the window length does not alter much the results of both test statistics.

Table 3 presents the results for the windowed testing. The third column in Table 3 shows the number of windows where the null of pure noise is rejected by the *C* statistic. In parentheses are the percentages of the total number of windows where this rejection occurs. For example, for the Bangkok S.E.T returns series, the null is rejected in 43 windows by the *C* statistic, which is equivalent to 41.35 per cent. Similarly, the fourth column provides the number and percentage of significant *H* windows. In those significant windows, the rejection of the null is due to either a significant *C* or *H*, or both. The statistics for total significant windows are provided in the final column of Table 3. The results clearly show that the number of significant windows for all the Asian stock markets is larger than the 1 per cent one would expect purely by chance, given the nominal threshold level of 0.01.[8] More importantly, the findings from this windowed test procedure demonstrate that the underlying returns series of all the Asian stock markets under investigation are characterised by transient epochs of dependencies. Specifically, these series can be described as a switching process, in which for a long period of time the returns follow a pure noise, interspersed with brief periods of highly significant linear and non-linear dependency structures. Hence, these underlying features violate the critical assumption of strict stationarity as required by the GARCH process.

[8] In this study, the threshold level has been set at 0.01. The level of significance is the bootstrapped thresholds that correspond to 0.01.

Table 3
Windowed Test Results for Binary Transformed Data

	Total Number of Windows	*Significant C Windows*	*Significant H Windows*	*Total Number of Significant C or H Windows*
Bangkok S.E.T.	104	43 (41.35%)	62 (59.62%)	62 (59.62%)
Hong Kong Hang-Seng	104	40 (38.46%)	49 (47.12%)	54 (51.92%)
Jakarta SE Composite	104	29 (27.88%)	31 (29.81%)	32 (30.77%)
Kuala Lumpur SE Composite	104	25 (24.04%)	34 (32.69%)	35 (33.65%)
Korea SE Composite	104	53 (50.96%)	66 (63.46%)	69 (66.35%)
Nikkei 225 Stock Average	104	32 (30.77%)	54 (51.92%)	56 (53.85%)
Philippines SE Composite	104	28 (26.92%)	49 (47.12%)	49 (47.12%)
Singapore Straits Times	104	22 (21.15%)	36 (34.62%)	36 (34.62%)

5. Implications and Concluding Remarks

The results from our econometric investigation using the Hinich portmanteau bicorrelation test show that the null of pure noise is strongly rejected by the *C* and *H* statistics. This implies that there are statistical structures present in the data of all the eight Asian stock markets that cannot be captured by the GARCH models. Further investigation via the windowed test procedure reveals that the violation of the covariance stationarity assumption as required by the GARCH process is due to the presence of transient epochs of dependencies in the underlying returns series. These episodic transient structures are also found in financial data of other stock markets (Ammermann and Patterson 2003) and foreign exchange markets (see, for example, Brooks and Hinich 1998; Brooks et al. 2000; Lim et al. 2003). It would be interesting to investigate whether the violation of covariance stationarity in the earlier studies of Pagan and Schwert (1990a, b) and Loretan and Phillips (1994) is due to similar structures present in the data.

One plausible explanation about the occurrence of these transient epochs of dependencies in all the Asian stock returns series is the market adjustment to unexpected shocks. In fact, it is possible to identify those specific unexpected events as the Hinich portmanteau bicorrelation test via the windowed test procedure permits a closer examination of the precise time periods during which these dependencies, linear and non-linear, are occurring. For instance, in the work of Ammermann and Patterson (2003), the linear dependencies that occurred episodically are found to be directly attributable to changes in the Taiwan Stock Exchange's price limits that were made during the years 1987 and 1988, even though the authors could not identify any clear-cut driving factor for the non-linear burst of dependencies. On the other hand, the non-linear dependency structures in Brooks et al. (2000) are found to be highly localised in time, and the authors referred to two important events that occurred during their sample period—widespread upsets in the currency markets and a change in US accounting procedures that affected US firms with businesses abroad. The explanation given is that when surprises hit the market, they usually generate a pattern of non-linear price movements relative to previous movements since the traders are unsure of how to react, and hence they respond slowly. Furthermore, as Schatzberg and Reiber (1992) pointed out, uninformed traders may delay their response to see how informed market participants behave because they do not have the resources to fully analyse the information or because the information is not reliable. Overreaction to bad news and underreaction to good news by investors and markets (see, for example, da Costa 1994; DeBondt and Thaler 1985, 1987), according to Antoniou et al. (1997), and Joe and Menyah (2003), could also trigger non-linear feedback mechanism. However, pursuing a detailed examination of the events that affect each of these Asian stock markets will be the subject of another article.

The episodic transient behaviour of dependencies detected in the returns series of all the Asian stock markets would pose a serious challenge to researchers attempting to model these series, especially after the popular GARCH models are found to be inadequate. Brooks and Hinich (1998) argued that even after modifications to the specifications of the GARCH models—such as those by Booth et al. (1994) to augment the GARCH equation with structural breaks in the mean, or those by Hamilton and Susmel (1994) to generalise the model for parameters to be drawn from one of several regimes—these dependencies could still not be captured. In fact, Hinich and Patterson (1995) conjectured that these features are responsible for the failure of non-linear models in providing superior out-of-sample forecasting performance, a question that has been raised by Diebold and Nason (1990) and further discussed by Ramsey (1996). Given the prevalence of these episodic transient features across all the Asian stock markets, coupled with the fact that their episodic occurrences are related to events specific to the particular market under study, no conclusion can be offered with confidence at this stage on models for replacement. However, Ammermann and Patterson (2003) are quite optimistic that some stable non-linear models could be developed to adequately describe these features in the data, though no recommendation has been offered.

On the other hand, since there is a growing trend of applying the GARCH models in these Asian stock markets, researchers and investors should take the findings in this study seriously. Specifically, the inadequacy of GARCH models for these Asian stock markets has strong implications for the pricing of stock index options contracts, portfolio selection, developing optimal hedging techniques and risk management.

Briefly, the risk-returns relationship is the central building block of most finance theory, for example the Capital Asset Pricing Model and Arbitrage Pricing Model. Related to this, volatility of returns is considered by researchers and investors to be a proxy for risk. Earlier studies have used unconditional standard deviation or variance as the measure of volatility and the limitations of such measures are well documented (see, for example, Jansen 1989; Pagan and Ullah 1988). Lately, the GARCH models have provided an alternative and useful measurement of volatility due to their ability to capture the clustering feature found in most financial time series data. Specifically, the GARCH estimate of time-varying volatility is a key input parameter for the pricing of options contracts and portfolio decisions. The significance of GARCH models is no less in constructing portfolio hedge ratios or modelling risk premia, in which the hedge ratios and risk premia are assumed to vary over time. Finally, most practical risk management decisions inherently rely on out-of-sample volatility forecasts, and GARCH models have always been the popular forecasting tools. However, the present findings indicate that the GARCH models might not be appropriate for the Asian stock markets as most of them are still categorised as emerging markets and have vastly different characteristics from those developed stock markets (see, for example, Bekaert and Harvey 1997).

To conclude, the present article provides empirical evidence against the adequacy of GARCH models in characterising the underlying process of all eight Asian stock markets under study. More importantly, the Hinich portmanteau bicorrelation test reveals that the violation of the covariance stationarity assumption is due to the presence of transient epochs of dependencies in the data. In addition to that, this study highlights the importance of conducting diagnostic checking, employing tests like the Hinich portmanteau bicorrelation test to determine the adequacy of the GARCH models before any further application in the financial world. Some possible extensions of the present work have been offered as well. To reiterate, it would be fruitful for future studies to determine the prevalence of these transient episodic structures in other financial markets, to conduct a detailed examination of the events that trigger these features in each of the Asian stock markets and finally to examine the candidates of models for such data.

Kian-Ping Lim (corresponding author) and **Venus Khim-Sen Liew** are at the Labuan School of International Business and Finance, Universiti Malaysia Sabah, Malaysia. E-mail: kianping@ums.edu.my
Melvin J. Hinich is at the Applied Research Laboratories, University of Texas at Austin.

References

Abhyankar, A.H., L.S. Copeland and **W. Wong** (1995), 'Nonlinear Dynamics in Real-Time Equity Market Indices: Evidence from the United Kingdom', *Economic Journal*, 105: 864–80.

Ammermann, P.A. and **D.M. Patterson** (2003), 'The Cross-Sectional and Cross-Temporal Universality of Nonlinear Serial Dependencies: Evidence from World Stock Indices and the Taiwan Stock Exchange', *Pacific-Basin Finance Journal*, 11: 175–95.

Antoniou, A., N. Ergul and **P. Holmes** (1997), 'Market Efficiency, Thin Trading and Non-Linear Behaviour: Evidence from an Emerging Market', *European Financial Management*, 3(2): 175–90.

Ashley, R.A. and **D.M. Patterson** (2001), 'Nonlinear Model Specification/Diagnostics: Insights from a Battery of Nonlinearity Tests', Working Paper, Virginia Polytechnic Institute and State University.

Barnett, W.A., A.R. Gallant, M.J. Hinich, J. Jungeilges, D. Kaplan and **M.J. Jensen** (1995), 'Robustness of Nonlinearity and Chaos Tests to Measurement Error, Inference Method, and Sample Size', *Journal of Economic Behavior and Organization*, 27: 301–20.

Barnett, W.A., A.R. Gallant, M.J. Hinich, J. Jungeilges, D. Kaplan and **M.J. Jensen** (1997), 'A Single-Blind Controlled Competition among Tests for Nonlinearity and Chaos', *Journal of Econometrics*, 82: 157–92.

Barnett, W.A. and **A. Serletis** (2000), 'Martingales, Nonlinearity, and Chaos', *Journal of Economic Dynamics and Control*, 24: 703–24.

Bekaert, G. and **C.R. Harvey** (1997), 'Emerging Equity Market Volatility', *Journal of Financial Economics*, 43: 29–77.

Bollerslev, T. (1986), 'Generalized Autoregressive Conditional Heteroskedasticity', *Journal of Econometrics*, 31: 307–27.

Bollerslev, T., R.Y. Chou and **K.F. Kroner** (1992), 'ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence', *Journal of Econometrics*, 52: 5–60.

Booth, G.G., T. Martikainen, S.K. Sarkar, I. Virtanen and **P. Yli-Olli** (1994), 'Nonlinear Dependence in Finnish Stock Returns', *European Journal of Operational Research*, 74: 273–83.

Brock, W.A., D.A. Hsieh and **B. LeBaron** (1991), *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*. Cambridge: MIT Press.

Brock, W.A., W.D. Dechert and **J.A. Scheinkman** (1987), 'A Test for Independence Based on the Correlation Dimension', Working Paper, University of Wisconsin, University of Houston and University of Chicago.

Brock, W.A., W.D. Dechert, J.A. Scheinkman and **B. LeBaron** (1996), 'A Test for Independence Based on the Correlation Dimension', *Econometric Reviews*, 15: 197–235.

Brooks, C. (1996), 'Testing for Non-Linearity in Daily Sterling Exchange Rates', *Applied Financial Economics*, 6: 307–17.

——— (1997), 'GARCH Modelling in Finance: A Review of the Software Options', *Economic Journal*, 107: 1271–76.

Brooks, C. and **M.J. Hinich** (1998), 'Episodic Nonstationarity in Exchange Rates', *Applied Economics Letters*, 5: 719–22.

Brooks, C., M.J. Hinich and **R. Molyneux** (2000), 'Episodic Nonlinear Event Detection: Political Epochs in Exchange Rates', in D. Richards (ed.), *Political Complexity: Political Epochs in Exchange Rates*, pp. 83–98. Michigan: Michigan University Press.

Brooks, C. and **S.M. Heravi** (1999), 'The Effect of Mis-Specified GARCH Filters on the Finite Sample Distribution of the BDS Test', *Computational Economics*, 13: 147–62.

Brooks, C., S.P. Burke and **G. Persand** (2001), 'Benchmarks and the Accuracy of GARCH Model Estimation', *International Journal of Forecasting*, 17: 45–56.

Choo, W.C., M.I. Ahmad and **M.Y. Abdullah** (1999), 'Performance of GARCH Models in Forecasting Stock Market Volatility', *Journal of Forecasting*, 18: 333–43.

da Costa, N.C.A. (1994), 'Overreaction in the Brazilian Stock Market', *Journal of Banking and Finance*, 18: 633–42.

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

——— (1987), 'Further Evidence on Investor Overreaction and Stock Market Seasonality', *Journal of Finance*, 42, 557–81.

Diebold, F.X. and **J.A. Nason** (1990), 'Nonparametric Exchange Rate Prediction?', *Journal of International Economics*, 28: 315–32.

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

——— (1995), *ARCH: Selected Readings*. Oxford: Oxford University Press.

Hall, S.G., D.K. Miles and **M.P. Taylor** (1989), 'Modelling Asset Prices with Time-Varying Betas', *The Manchester School*, 57(4): 340–56.

Hamilton, J.D. and **R. Susmel** (1994), 'Autoregressive Conditional Heteroskedasticity and Changes in Regime', *Journal of Econometrics*, 64: 307–33.

Hentschel, L. (1995), 'All in the Family: Nesting Symmetric and Asymmetric GARCH Models,' *Journal of Financial Economics*, 39: 71–104.

Hinich, M.J. (1982), 'Testing for Gaussianity and Linearity of a Stationary Time Series', *Journal of Time Series Analysis*, 3: 169–76.

——— (1996), 'Testing for Dependence in the Input to a Linear Time Series Model', *Journal of Nonparametric Statistics*, 6: 205–21.

Hinich, M.J. and **D.M. Patterson** (1995), 'Detecting Epochs of Transient Dependence in White Noise', Mimeo, University of Texas at Austin.

Hooy, C.W. and **H.B. Tan** (2002), 'Volatility Spillover Effects among Major Asia Pacific Equity Markets', in M. Azali, M.B. Yusoff, M.S. Habibullah, S.A. Mansor, R. Mustafa, S. Mohamad, Z. Yusop, A. Radam, A. Hassan and S. Pillai (eds), *Proceedings: Asia Pacific Economics and Business Conference 2002—The Impact of Global Competitive Environment on Asian Economics, Business and Society (Volume 2),* pp. 911–18. Serdang: Universiti Putra Malaysia Press.

Hsieh, D.A. (1989), 'Testing for Nonlinear Dependence in Daily Foreign Exchange Rates', *Journal of Business*, 62: 339–68.

Hsieh, D.A. (1991), 'Chaos and Nonlinear Dynamics: Application to Financial Markets', *Journal of Finance*, 46: 1839–77.

Jansen, D.W. (1989), 'Does Inflation Uncertainty Affect Output Growth? Further Evidence', *The Federal Reserve Bank of St. Louis Review*, July–August: 43–54.

Joe, A.K. and **K. Menyah** (2003), 'Return Predictability in African Stock Markets', *Review of Financial Economics*, 12: 247–70.

Krager, H. and **P. Kugler** (1993), 'Non-Linearities in Foreign Exchange Markets: A Different Perspective', *Journal of International Money and Finance*, 12: 195–208.

LeBaron, B. (1997), 'A Fast Algorithm for the BDS Statistic', *Studies in Nonlinear Dynamics and Econometrics*, 2(2): 53–59.

Lim, K.P. (2001), 'Non-Linear Dependence in ASEAN-5 Foreign Exchange Rates: An Insight from a Battery of Non-Linearity Tests'. Unpublished MSc thesis, Universiti Putra Malaysia, Malaysia.

Lim, K.P., M.J. Hinich and **V.K.S. Liew** (2003), 'Episodic Non-Linearity and Non-Stationarity in ASEAN Exchange Rates Returns Series', *Labuan Bulletin of International Business and Finance*, 1: 79–93.

Loretan, M. and **P.C.B. Phillips** (1994), 'Testing the Covariance Stationarity of Heavy-Tailed Time Series', *Journal of Empirical Finance*, 1: 211–48.

McCullough, B.D. and **C.G. Renfro** (1999), 'Benchmarks and Software Standards: A Case Study of GARCH Procedures', *Journal of Economic and Social Measurement*, 25: 59–71.

McMillan, D.G. and **A.E.H. Speight** (2001), 'Nonlinearities in the Black Market Zloty-Dollar Exchange Rate: Some Further Evidence', *Applied Financial Economics*, 11: 209–20.

Opong, K.K., G. Mulholland, A.F. Fox and **K. Farahmand** (1999), 'The Behavior of Some UK Equity Indices: An Application of Hurst and BDS Tests', *Journal of Empirical Finance*, 6: 267–82.

Pagan, A.R. (1996), 'The Econometrics of Financial Markets', *Journal of Empirical Finance*, 3: 15–102.

Pagan, A.R. and **A. Ullah** (1988), 'The Econometric Analysis Models with Risk Terms', *Journal of Applied Econometrics*, 3: 87–105.

Pagan, A.R. and **A.W. Schwert** (1990a), 'Alternative Models for Conditional Stock Volatility', *Journal of Econometrics*, 45: 267–90.

——— (1990b), 'Testing for Covariance Stationarity in Stock Market Data', *Economic Letters*, 33: 165–70.

Psaradakis, Z. (2001) '*p*-Value Adjustments for Multiple Tests for Nonlinearity', *Studies in Nonlinear Dynamics and Econometrics*, 4(3): 95–100.

Ramsey, J.B. (1996), 'If Nonlinear Models Cannot Forecast, What Use are They?', *Studies in Nonlinear Dynamics and Econometrics*, 1(2): 65–86.

Schatzberg, J.D. and **R.R. Reiber** (1992), 'Extreme Negative Information and the Market Adjustment Process: The Case of Corporate Bankruptcy', *Quarterly Journal of Business and Economics*, 31: 3–21.

Tan, J.H. (2001), 'Time-Varying Volatility of the Malaysian Stock Index Futures Market: Persistency, Informational Shocks, Price Levels and Selective Capital Controls', *Banker's Journal*, 119: 12–21.