Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 15 (2008) 518 – 532
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

Are Asian stock markets efficient? Evidence from new multiple variance ratio tests ☆

Jae H. Kim [a,*], Abul Shamsuddin [b]

[a] Department of Econometrics and Business Statistics, Monash University, Caulfield East, VIC 3145, Australia
[b] Newcastle Graduate School of Business, University of Newcastle, Callaghan, NSW 2308, Australia

Received 6 October 2006; received in revised form 14 June 2007; accepted 2 July 2007
Available online 14 July 2007

Abstract

This paper tests for the martingale hypothesis in the stock prices of a group of Asian markets. We use new multiple variance ratio tests based on the wild bootstrap and signs. These are non-parametric finite sample tests, which do not rely on large sample theories for statistical inference. This paper also presents Monte Carlo results that these non-parametric tests show superior small sample properties to those of the conventional Chow–Denning test. Both weekly and daily data from 1990 are considered, while moving sub-sample windows are used for the latter to control the sensitivity of the results to a particular sample period. It is found that the Hong Kong, Japanese, Korean and Taiwanese markets have been efficient in the weak-form. The markets of Indonesia, Malaysia and Philippines have shown no sign of market efficiency, despite financial liberalization measures implemented since the eighties. We have also found evidence that the Singaporean and Thai markets have become efficient after the Asian crisis. In general, the results point toward the notion that the pricing efficiency of a market depends on the level of equity market development as well as the regulatory framework conducive of transparent corporate governance.

JEL classification: G14; G15
Keywords: Martingale hypothesis; Stock market efficiency; Wild bootstrap; Joint; Sign test

1. Introduction

Finance practitioners often operate under the presumption that stock prices do not follow a random walk, while most finance academics believe in the opposite premise. Most academics still believe in an efficient capital market, where stock prices at any time fully reflect all available and relevant information.[1] Therefore, given only past price and return

☆ Earlier versions of this paper were presented at the Australian Econometrics Society Meeting 2004, Melbourne, Australia; International Forum on Monetary and Financial Cooperation for Asia, Seoul National University, Korea in 2004; and seminars at La Trobe University and Monash University. We would like to thank Graham Smith and Param Silvapulle for useful comments, and the referees for their comments that have substantially improved the paper. All calculations are conducted using the GAUSS, and the GAUSS codes are available from the authors on request.
* Corresponding author. Tel.: +61 3 99031596; fax: +61 3 99032007.
E-mail address: Jae.Kim@BusEco.monash.edu.au (J.H. Kim).
[1] This belief was shaken by the technology bubble of the late 1990s. See Stangle (2005) for a synopsis of the debate over the efficient market hypothesis between Burton Malkiel and Sendhil Mullainathan, two leading proponents of market efficiency and behavioral finance, respectively.

0927-5398/$ - see front matter
doi:10.1016/j.jempfin.2007.07.001

data, the current price is the best predictor of the future price, and the price change or return is expected to be zero. This is the essence of the weak-form efficient market hypothesis, which implies a random walk. While a random walk requires its increments to be identically and independently distributed, a martingale allows for uncorrelated increments with a general form of heteroskedasticity. Hence, the martingale model is a generalized version of the random walk, better suited to asset prices with conditionally heteroskedastic increments. If an asset price follows a martingale, then its return is purely unpredictable and investors are unable to make abnormal returns over time. Hence, the question as to whether an asset price follows a martingale has strong implications to the efficient market hypothesis in the weak-form.

The central aim of this paper is to test for the martingale property of a number of Asian stock markets, which are the markets of considerable interest to both finance practitioners and academics. This region includes dynamic economies with high growth and increasing stock market volumes, and its major players consist of developed as well as emerging markets. The emerging markets in this region also have been exposed to financial liberalization since the late 1980's or early 1990's, and its gradual influence on the stock market efficiency to date is of considerable interest (see Kawakatsu and Morey, 1999). Moreover, the Asian crisis in 1997 heavily affected many countries in this region, and its impact on stock market efficiency should be examined. Following the Asian crisis, the OECD and the World Bank have been encouraging the policy makers, regulators and market participants in Asian countries to adopt international standard and practices for financial and non-financial disclosure (OECD, 2003). It is widely believed that, since the Asian crisis, the structure and operations of Asian stock markets have been markedly strengthened, which in turn may have improved the informational efficiency of these markets. As we shall see in the next section, there have been many studies that tested efficiency of Asian stock markets. However, the results are overall mixed and scattered over numerous studies that use different sample periods, methods and data frequencies. There is a need for a comprehensive study, paying attention to the effects of the financial liberalization and Asian crisis.

As a means of testing for market efficiency, the variance ratio test is used in this paper, which has become the most popular tool since the work of Lo and MacKinlay (1988). On the methodological aspect, we have identified the following deficiencies in past studies. First, many stock markets frequently experience structural changes or outliers, due to the events such as financial catastrophe and changes in regulatory regime. The structural change can alter the nature of market efficiency, while influential outliers adversely affect the performance of statistical tests. Past studies on the Asian markets did not attempt to separate the effects of (potential) structural changes or influential outliers that may have distorted the test outcomes on market efficiency. Second, most past studies adopted asymptotic tests that have deficient small samples properties. For example, the conventional variance ratio tests can possess the sampling distribution far different from their asymptotic distribution in small sample (see Lo and MacKinlay, 1989). This can lead to severe size distortions or low power in small samples, resulting in misleading statistical inference.

This study contributes to this ongoing debate over the market efficiency of Asian stock markets with the following distinct features. First, we pay attention to the effects of the financial liberalization and Asian crisis using stock market indices over a long period. Second, this paper conducts the variance ratio test using a moving sub-sample window. This is to accommodate the dynamic nature of Asian stock markets, and to obtain inferential outcomes robust to possible structural changes or influential outliers.[2] Third, we employ non-parametric variance ratio tests that do not rely on asymptotic approximations and possess desirable small sample properties.

We prefer the use of multiple variance ratio tests in order to control the size of the joint test. This is because the martingale hypothesis is a joint hypothesis in the context of the variance ratio test, and the use of individual test for a joint hypothesis will necessarily induce size distortion (see Savin, 1984).[3] The multiple variance ratio tests adopted in this paper are the wild bootstrap test of Kim (2006) and a multiple version of Wright's (2000) sign test, in addition to the conventional Chow and Denning (1993) test. We report the Monte Carlo results that these new non-parametric tests have superior small sample properties to those of the Chow–Denning test.

We consider the stock markets of nine Asian countries including Hong Kong, Indonesia, Japan, Korea, Malaysia, Philippines, Singapore, Taiwan and Thailand using both daily and weekly stock price indices from 1990 to 2005. The sample includes stock markets with varying degree of development, enabling us to find out whether the weak-form market efficiency is contingent upon the level of equity market development. It can be argued that developed equity

[2] This technique also has been adopted in recent studies such as Yilmaz (2003) and Belaire-Franch and Opong (2005a) to examine the martingale property of exchange rates and FTSE indices, respectively.

[3] In the same context, Belaire-Franch and Opong (2005b) considered size adjustment for Wright's (2000) individual tests for a joint hypothesis, while Fong et al. (1997) and Yilmaz (2003) used the multiple tests.

market facilitates efficient transmission information, which may lead to pricing efficiency. The paper finds that market efficiency varies with the level of equity market development, which is measured typically by the FTSE country classification.[4] The FTSE classifies all countries into developed, advanced emerging and secondary emerging. Our sample includes three developed (Hong Kong, Japan, and Singapore), two advanced emerging (Korea and Taiwan) and four secondary emerging (Indonesia, Malaysia, Philippines and Thailand) countries. Overall, we have found that the developed or advanced emerging markets show weak-form efficiency, while the secondary emerging markets are found to be inefficient. Some markets have become efficient after the Asian crisis in 1997. Despite financial market liberalization, some emerging markets have shown little sign of market efficiency.

The structure of the paper is as follows. In the next section, a review of past studies on stock market efficiency is given, paying attention to the case of Asian stock markets. Section 3 provides the details and small sample properties of the variance ratio test to be used in this paper. Section 4 presents the data and computation details, and Section 5 the empirical results. The conclusion is drawn in Section 6.

2. A review of past studies on stock market efficiency

Testing for the random walk model has been a subject of much attention in the empirical finance literature, since it is a requirement of the weak-form efficient market hypothesis. This model has two testable implications. First, stock returns are not predictable using past price information. Second, the variance of return is linearly associated with holding period. The former has been tested in the literature by examining the serial correlation structure of stock returns, while the latter using the variance ratio test proposed by Lo and MacKinlay (1988).

The first generation of studies on weak-form market efficiency primarily examined serial correlation of price changes. In particular, the pioneering work of Kendall (1953) and Fama (1965) on the random walk hypothesis observed no serial correlation in returns, providing evidence in support of the weak-form efficient market hypothesis. Fama's (1970) comprehensive review of the early literature on market efficiency also provided overwhelming support for the efficient market hypothesis. The subsequent empirical work, however, has provided mixed results.

Some ground-breaking studies on the random walk hypothesis emerged in the 1980s. Unlike the first generation analyses, most of these studies with superior econometric methodology provided evidence against the random walk hypothesis. This evidence of market inefficiency was explained in terms of market overshooting (Dedondt and Thaler, 1985), noise trading (Black, 1986), time-varying risk and returns (Lo and MacKinlay, 1988) and fads (Lehmann, 1990). Two major studies in the cohort of the second generation studies are Poterba and Summers (1988) and Fama and French (1988). Among competing statistical techniques of testing for the random walk hypothesis, Poterba and Summers (1988) preferred the variance ratio test. Investigating monthly returns for the US and 17 other stock markets, they observed negative serial correlation in the long horizon and positive serial correlation in the short horizon, which provided evidence in support of the mean reversion property, contrary to the random walk hypothesis. They concluded that serial correlation was attributed to noise trading, not fundamental factors. Their study highlighted the possibility that the finding on the random walk hypothesis might be sensitive to sampling intervals. This issue was examined comprehensively by Fama and French (1988). They used an adjusted OLS model to estimate autocorrelation coefficients for 1 to 10 years of real returns. In general, their results indicated a U-shape pattern in autocorrelation coefficients with respect to investment horizons and provided some evidence against the random walk hypothesis. However, these results were primarily driven by data from pre-World War II period. The evidence of the U-shape and large autocorrelations was much weaker in the post-World War II period.

Following the works of Poterba and Summers (1988) and Fama and French (1988), many researchers attempted to use alternative tests for a random walk hypothesis and challenged the earlier findings. For example, Lo (1991) used modified rescaled range test (R/S test) to demonstrate absence of long-term memory in stock prices and support the random walk hypothesis. Further studies also reached the same conclusion using the joint test of modified R/S test and

[4] It is worth noting that a thorough examination of the level of equity market development and its empirical relationship with the degree of market efficiency is beyond the scope of the study. However, as an attempt to appreciate the rationale behind our key empirical findings on the weak form efficiency, we try to detect a possible link between the level of equity market development and the degree of informational efficiency in Section 5 of this paper. To this end we merely rely on a widely used indicator of equity market development — the FTSE classification of equity markets, which takes into account four broad criteria: economic size, wealth, quality of markets and depth and breadth of markets. For further details about the FTSE classification see http://www.ftse.com/Indices/Country_Classification/index.jsp.

the variance ratio test (see Chow and Denning, 1993; Richardson, 1993). Using the variance ratio test and the modified rescaled range test, Coggin (1998) and Berneburg (2004) were also unable to reject the random walk hypothesis for the US and European equity indexes, respectively.

A large number of studies emerged over the last decade to investigate the efficiency of Asian stock markets. Table 1 presents a concise summary of the findings of the selected studies, indicating the methodologies used and the types of data employed. Generally, the results were sensitive to the sampling frequency. The hypothesis of weak-form efficiency could be rejected using daily data, but not using lower frequency data (weekly, monthly, quarterly, etc.). This finding is consistent with the notion of slow price adjustment in response to a shock, which might be attributed to regulatory constraints such as the system of stock price limits and the extent of domestic financial market liberalization. Overall, the past studies on Asian stock market efficiency provide mixed results, although they tend to indicate that the emerging Asian markets are not efficient in the weak-form.

However, the results are scattered over numerous studies, and many of them are too outdated to reflect the impact of the Asian crisis and gradual effects of financial liberalization.

This paper conducts a comprehensive analysis of the issue for nine Asian stock markets using updated data. We prefer weekly and daily data in this paper, because a large sample size is required to ensure reasonably high power of the multiple variance ratio tests to be used in this paper (See Section 3.4). Chow and Denning (1993) also provided Monte Carlo evidence that their multiple variance ratio tests require the sample size of at least 256 for reasonable power properties. The use of monthly data would be interesting as they are less noisy than daily or weekly data. We leave this possibility for future research, due mainly to data availability, because the moving sub-sample window is operational with a large number of observations.

3. The variance ratio tests

The variance ratio (VR) test is based on the property that, if a time series of an asset return is a purely random, the variance of k-period return is k times the variance of one-period return. Hence, the VR, defined as the ratio of $1/k$ times the variance of k-period return to that of one-period return, should be equal to one for all values of k. As mentioned earlier, the multiple VR tests are preferred in this paper, whose null hypothesis is that a set of VR's over a number of

Table 1
Selected studies on weak-form efficiency of Asian markets

Studies	Data and methodology	Markets	RWH hypothesis
Mun and Kee (1994)	DW, 1982–1991 VR, Spectral shape test	HK, KOR, THA, MAL, TAW	Rejected for all countries with daily data Rejected only MAL and THA with weekly data
Ayadi and Pyun (1994)	D, W, M, 2M, Q, 1984–1998 VR	KOR	Rejected with daily data Not rejected with lower frequency data
Huang (1995)	W, 1988–1992 VR, ADF	HK, IND, KOR, JAP PHI, SIN, THA, TAW	Rejected for KOR, MAL, HK, SIN, THA
Kawakatsu and Morey (1999)	M, 1976–1997 VR, MVR, ADF	KOR, MAL, PHI, TAW, THA	Find little evidence that financial market liberalization improves efficiency
Ryoo and Smith (2002)	D, 1988–1998 MVR	KOR	The market approaches RW as price limits are relaxed
Chang and Ting (2000)	W, M, Q, Y, 1971–1996 VR	TAW	Rejected with daily data Not rejected with lower frequency data
Peng et al. (2004)	W, 1993–2002 Chaos theory	TAW	Rejected
Lima and Tabak (2004)	D, 1992–2000 MVR	China, HK, SIN	Rejected for China and SIN
Lee et al. (2001)	D, 1990–1997 VR, ADF	China	Rejected
Hoque et al. (in press)	W, 1990–2004 VR, MVR, W	HK, IND, KOR, MAL PHI, SIN, THA, TAW	Rejected except for KOR and TAW

D: daily, W: weekly, M: monthly, Q: quarterly, Y: annual; VR: Lo–MacKinlay variance ratio test; MVR: Chow–Denning multiple variance ratio test, W: Wright's sign test; ADF: Augmented Dickey–Fuller Test; HK: Hong Kong; JAP: Japan, KOR: Korea, TAW: Taiwan, IND: Indonesia, PHI: Philippines; MAL: Malaysia, THA: Thailand, SIN: Singapore; All studies listed in this table used aggregate stock market indices to test market efficiency.

holding periods is jointly equal to one. In this section, we present a review of multiple VR tests applicable to test for the martingale hypothesis.[5]

Let x_t be an asset return at time t, where $t=1,\ldots, T$. It is assumed that x_t is a realization of the underlying stochastic process X_t, which follows a martingale difference sequence. This means that X_t's are serially uncorrelated, but are allowed to be conditionally or unconditionally heteroskedastic. Following the formula given by Wright (2000), the VR statistic be written as

$$\mathrm{VR}(x;k) = \left\{\frac{1}{Tk}\sum_{t=k}^{T}(x_t + x_{t-1} + \ldots + x_{t-k+1} - k\hat{\mu})^2\right\} \div \left\{\frac{1}{T}\sum_{t=1}^{T}(x_t - \hat{\mu})^2\right\}, \tag{1}$$

where $\hat{\mu} = T^{-1}\sum_{t=1}^{T} x_t$. This is an estimator for the unknown population VR, denoted as $V(k)$, which is the ratio of $1/k$ times the variance of k-period return to the variance of one-period return.

Lo and MacKinlay (1988, Assumption H^*) assumed the following conditions for X_t which are restated as below:

1. For all t, $E(X_t)=0$ and $E(X_t X_{t-k})=0$ for any $k\neq 0$;
2. X_t is a ϕ-mixing sequence with coefficient $\phi(m)$ of size $r/(2r-1)$ or is α-mixing[6] with coefficients $\alpha(m)$ of $r/(r-1)$, where $r>1$, for all t and for any $k\geq 0$, there exists some $\delta>0$ for which $E|X_t X_{t-k}|^{2(r+\delta)}<\infty$;
3. $\lim_{T\to\infty}\frac{1}{T}\sum E(X_t^2)<\infty$; and
4. For all t, $E(X_t X_{t-j} X_t X_{t-k})=0$ for all non-zero j and k where $j\neq k$.

According to Lo and MacKinlay (1988; p.49), the return sequence X_t satisfying the above conditions is serially uncorrelated with a general form of heteroskedasticity including deterministic change of variance or ARCH-type conditional variance. Lo and MacKinlay (1988) showed that, under the null hypothesis that $V(k)=1$,

$$M(x;k) = (\mathrm{VR}(x;k) - 1)\left(\sum_{j=1}^{k-1}\left[\frac{2(k-j)}{k}\right]^2 \delta_j\right)^{-1/2} \tag{2}$$

follows the standard normal distribution asymptotically under Assumption H^*, where

$$\delta_j = \left\{\sum_{t=j+1}^{T}(x_t - \hat{\mu})^2(x_{t-j} - \hat{\mu})^2\right\} \div \left\{\left[\sum_{t=1}^{T}(x_t - \hat{\mu})^2\right]^2\right\}.$$

3.1. Chow–Denning test

Chow and Denning (1993) proposed a multiple VR test for the joint null hypothesis $V(k_i)=1$ for $i=1,\ldots, l$ against the alternative hypothesis that $V(k_i)\neq 1$ for some holding period k_i. The test statistic can be written as

$$\mathrm{MV} = \max_{1\leq i\leq l}|M(x;k_i)|. \tag{3}$$

This is based on the idea that the decision regarding the null hypothesis can be made based on the maximum absolute value of individual VR statistics. Under Assumption H^* of Lo and MacKinlay (1988), the statistic follows the studentized maximum modulus distribution with l and T degrees of freedom, whose critical values are tabulated in Stoline and Ury (1979). When T is large, the null hypothesis is rejected at α level of significance if the MV statistic is greater than the $[1-(\alpha^*/2)]$th percentile of the standard normal distribution where $\alpha^*=1-(1-\alpha)^{1/l}$.

[5] The VR tests applicable only to iid or conditionally homoskedastic return time series, such as the iid version of the Lo–MacKinlay test, rank test of Wright (2000) and Wald test of Richardson and Smith (1991), are not included. The sub-sampling test of Whang and Kim (2003) is also excluded because it shows small sample properties far inferior to the wild bootstrap test under the sample size of interest in this paper (see Kim, 2006).

[6] Further details of mixing appear in Chapter 3 of White (1999).

3.2. Wild Bootstrap test

The bootstrap is a resampling method that approximates the sampling distribution of a test statistic (Efron, 1979).[7] To implement the bootstrap to time series data with unknown forms of heteroskedasticty, the wild bootstrap of Mammen (1993) is adopted in this paper. Following Kim (2006), the wild bootstrap for the Chow–Denning test given in Eq. (3) can be conducted in three stages as below:

(i) Form a bootstrap sample of T observations $x_t^* = \eta_t x_t$ ($t=1,\ldots, T$) where η_t is a random sequence with zero mean and unit variance.
(ii) Calculate MV*, which is the MV statistic in Eq. (3) obtained from the bootstrap sample generated in stage (i).
(iii) Repeat (i) and (ii) sufficiently many, say m, times to form a bootstrap distribution of the test statistic $\{MV^*(j)\}_{j=1}^m$.

The bootstrap distribution $\{MV^*(j)\}_{j=1}^m$ is used to approximate the sampling distribution of the MV statistic. The p-value of the test is estimated as the proportion of $\{MV^*(j)\}_{j=1}^m$ greater than the MV statistic calculated from the original data.

Under Assumption H^* of Lo and MacKinlay (1988), MV* has the same limiting distribution as MV given in Eq. (3), according to Kim (2006). In implementing the wild bootstrap test, a specific form of η_t should be chosen. In this paper, we follow Kim (2006) in using the standard normal distribution for η_t since Kim (2006) reported that other choices provided qualitatively similar small sample results.

3.3. Joint Sign test

Consider $s_t = u(x_t,0)$, where $u(x_t,0)=1$ if $x_t > 0$, and 0 otherwise. Under the assumption that x_t is generated from a martingale difference sequence (Assumptions A1 and A2 of Wright, 2000), s_t is an i.i.d. sequence with zero mean and unit variance, which takes the value 1 and − 1 with equal probability of 0.5. According to Wright (2000), the Assumptions A1 and A2 are neither strictly weaker nor strictly stronger than Assumptions H^* of Lo and MacKinlay (1988), while the former include popular conditionally heteroskedastic models such as the GARCH and stochastic volatility.

Wright (2000) proposed a sign-based test for an individual null hypothesis $V(k)=1$ against the alternative hypothesis that $V(k)\neq 1$. The test statistic can be written as

$$S_1(k) = \left(\frac{(Tk)^{-1} \sum_{t=k}^{T} (s_t + s_{t-1} + \ldots + s_{t-k+1})^2}{T^{-1} \sum_{t=1}^{T} s_t^2} - 1 \right) \left(\frac{2(2k-1)(k-1)}{3kT} \right)^{-1/2}. \tag{4}$$

The exact sampling distribution of Eq. (4) is given in Proposition 2 of Wright (2000), and the critical values of the test can be obtained by simulating its sampling distribution. To test for the joint null hypothesis that $V(k_i)=1$ for $i=1, \ldots, l$ against the alternative hypothesis that $V(k_i)\neq 1$ for some i, we propose a multiple test of the following form:

$$JS = \max_{1\le i\le l} |S_1(k_i)|, \tag{5}$$

similarly to the Chow and Denning (1993) test. The JS statistic also has an exact sampling distribution, and its critical values can be obtained by simulation in a similar way as that of $S_1(k)$ given in Eq. (4). The null hypothesis is rejected when the observed JS statistic is greater than the critical value.[8] We also note that Belaire-Franch and Contreras (2004) independently proposed a joint sign test, in a similar fashion to the one proposed in this paper.[9]

[7] It has been applied widely in many areas of time series econometrics and statistics, and often found to perform better than asymptotic approximations in small samples (see, Li and Maddala, 1996; Berkowitz and Killian, 2000).

[8] As an example, for $k_i=\{2, 5, 10, 20, 40\}$, the 5% and 10% critical values are respectively 2.40 and 2.10 when $T=500$; and 2.41 and 2.10 when $T=1000$ (based on 10000 Monte Carlo trials).

[9] We would like to thank a referee for bring this work to our attention.

3.4. Small sample properties of joint VR tests

Kim (2006) examined small sample properties of the wild bootstrap tests. It is found that the MV* test has correct size, improving the size distortions of the MV test. Kim (2006) also reported that the MV* test has higher power than the MV test. As the JS test is an exact test, it also shows no size distortion in small samples. However, since its power properties are unknown, it is instructive to evaluate the power properties of the JS test, making comparison to those of the MV* and MV tests. In this section, a small scale Monte Carlo experiment is conducted for this purpose, following the experimental design of Kim (2006).

We consider AR(1) and long memory (LM) models, which can be written respectively as

$$X_t = 0.1X_{t-1} + u_t; \text{ and } (1-B)^{0.1}X_t = u_t,$$

where B is the backshift operator. These models represent the return time series with the types of serial correlation widely observed in practice (see, for example, Wright, 2000; Kim, 2006). The AR(1) coefficient and long memory parameter values are set to 0.1 to simulate the return sequences with a considerably low degree of serial correlation, which are difficult to distinguish from an uncorrelated sequence. We consider two forms of disturbance term

$$u_t = \sqrt{h_t}\varepsilon_t; h_t = 0.5 + 0.75h_{t-1} + 0.1\varepsilon_{t-1}^2(\text{GARCH}(1,1)); \text{ and}$$

$$u_t = \exp(0.5h_t)\varepsilon_t; h_t = 0.95h_{t-1} + \varepsilon_t(\text{stochastic volatility, SV}).$$

For the error term ε_t, we have considered the standard normal as well as Student's t distribution with three degrees of freedom. This gives GARCH(1,1) disturbance terms with the standard normal and Student's t errors (denoted as GARCH1 and GARCH2 respectively), and similarly for SV disturbance terms (denoted as SV1 and SV2). These disturbance terms represent those widely observed from daily and weekly financial returns, which were also used in the past simulation studies such as Lo and MacKinlay (1989) and Wright (2000). The probability of rejection (or power of the test) is calculated over 1000 Monte Carlo trials when the level of significance is 5%. The values of holding periods

Table 2
Probability of rejection from alternative multiple VR tests

	GARCH1	GARCH2	SV1	SV2
AR(1) Model				
T=500				
MV	0.321	0.317	0.189	0.228
MV*	0.403	0.415	0.291	0.307
JS	0.207	0.231	0.221	0.370
T=1000				
MV	0.641	0.623	0.368	0.403
MV*	0.703	0.692	0.489	0.522
JS	0.378	0.362	0.408	0.597
Long memory				
T=500				
MV	0.588	0.590	0.469	0.483
MV*	0.670	0.689	0.521	0.584
JS	0.489	0.482	0.541	0.769
T=1000				
MV	0.882	0.892	0.725	0.768
MV*	0.937	0.932	0.765	0.797
JS	0.719	0.736	0.807	0.943

MV: Chow–Denning test, MV*: wild bootstrap test, JS: joint sign test; GARCH1: GARCH(1,1) with standard normal distribution.
GARCH2: GARCH(1,1) with Student-t distribution with three degrees of freedom.
SV1: Stochastic volatility with standard normal distribution.
SV2: Stochastic volatility with Student-t distribution with three degrees of freedom.

(k_i's) considered are (2, 5, 10, 20, 40), and the sample sizes considered are 500 and 1000. The number of bootstrap iterations m for the wild bootstrap test is set to 1000.

The results are reported in Table 2. The MV* test shows higher power than the JS test under GARCH1 and GARCH2. The JS test, on the other hand, has higher power than the MV* test under SV1 and SV2, except for the case of AR(1) model with SV1 disturbance term. Hence, we observe a strong tendency that the MV* test perform better than the JS test under GARCH(1,1) disturbance terms, while the JS test in general outperform the MV* test under SV disturbance terms. The results indicate that the MV* and JS tests are good alternatives in testing for the martingale property of a financial return time series. It is also evident that these VR tests show high power when the sample size is as large as 500 and the power rises considerably as the sample size increases from 500 to 1000, a feature also evident from Monte Carlo results reported in Chow and Denning (1993). Although the details are not reported for simplicity, the simulations were also conducted with much longer holding periods. We found that both tests maintain desirable power properties with longer holding periods, although there is tendency that the power of the tests gets, to a degree, lower with longer holding periods. These additional Monte Carlo results are available from the corresponding author on request.

4. Data and computation details

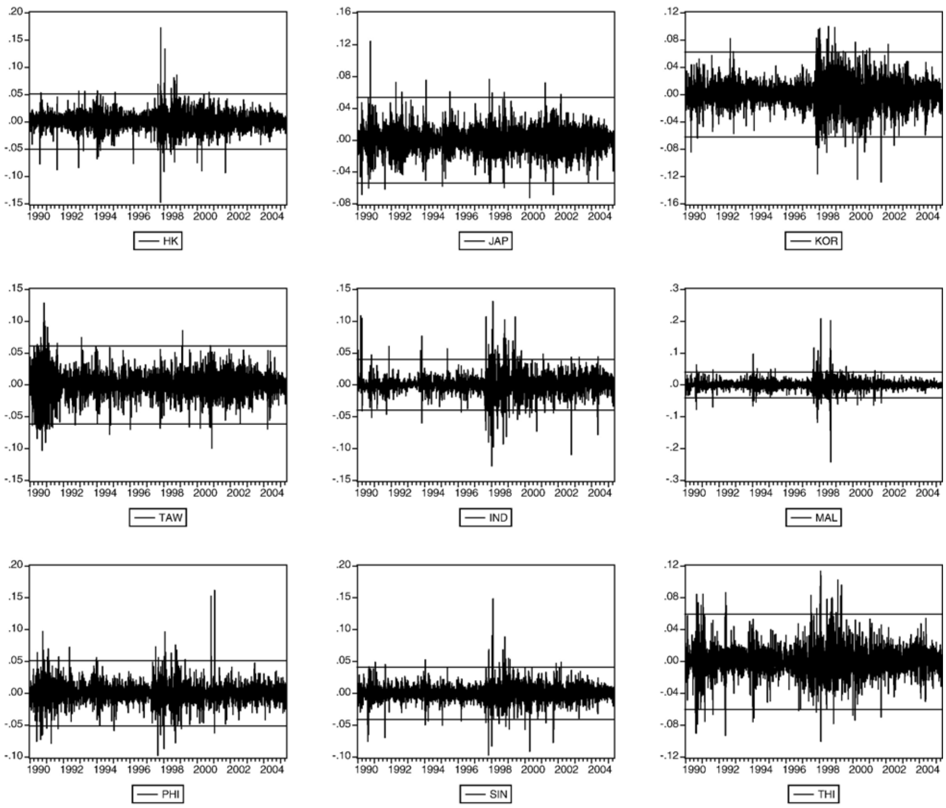
We consider the stock price indices for the markets from the following countries: Hong Kong, Japan, Korea, Indonesia, Malaysia, Philippines, Singapore, Taiwan and Thailand. The log returns (nominal) are derived from the stock price indices denominated in local currency; i.e. $x_t = \log(P_t/P_{t-1})$ where P_t is the value of the index at time t and log is the natural logarithm. All data are obtained from Datastream. Table 3 provides the index descriptions and Datastream codes, with detailed information about each index.

The data covers the period from 1 January 1990 to 29 April 2005, comprising 4000 daily and 800 weekly observations. The weekly data are associated with Wednesdays, but Tuesday data points are used when Wednesday is a holiday. According to the descriptive statistics and the ARCH LM test, all log return series are highly non-normal with a strong degree of conditional heteroskedasticity, as might be expected from weekly and daily stock returns. The details of descriptive statistics are not reported for simplicity. Fig. 1 presents the time plots of daily log returns for all markets. The horizontal lines in each graph indicate (Q1 − 3IQR, Q3 + 3IQR), where Q1 is the first quartile, Q3 is the third quartile, and IQR is the inter-quartile range (Q3 − Q1). This is a popular criterion to detect extreme outliers based on the

Table 3
Details of the stock price indices used

Markets	Index name	Datastream code
Hong Kong (HK)	Hang Seng	HNGKNGI
Japan (JAP)	Nikkei 225 Stock Average	JAPDOWA
Korea (KOR)	Korea SE Composite (KOSPI)	KORCOMP
Taiwan (TAW)	Taiwan SE (TWSE) Weighted	TAIWGHT
Indonesia (IND)	Jakarta SE Composite	JAKCOMP
Malaysia (MAL)	KLCI Composite	KLPCOMP
Philippines (PHI)	Philippine SE I(PSEi)	PSECOMP
Singapore (SIN)	Singapore Straits Times (New)	SNGPORI
Thailand (THI)	Bangkok S.E.T.	BNGKSET

Notes: The Hang Seng price index is a capitalization-weighted index of a portfolio of 33 largest companies listed on the Hong Kong Exchanges and Clearing Limited (HKEx). It represents about 70% of market capitalization of the HKEx. The NIKKEI 225 is a price-weighted index of 225 largest companies trading on the first section of the Tokyo Stock Exchange. The KOSPI is a market capitalization based index comprising all companies listed on the Korea Stock Exchange. The TWSE Index and the Bangkok SET Index are capitalization-weighted indexes of all listed stocks traded on the Taiwan Stock Exchange and the Stock Exchange of Thailand, respectively. For the Singapore Stock Exchange, the Straits Times Index is used, which is a modified value-weighted index consisting 55 of the largest companies listed on the Singapore Stock Exchange. For the Malaysian stock market, the Kuala Lumpur Composite Index is used. It is a market capitalization based index, including 100 of the largest companies listed on the Kuala Lumpur Stock Exchange. The Philippines SE Composite (complied by Datastream) is a market capitalization weighted index comprising 30 of the largest companies listed on the Philippines Stock Exchange. The Jakarta Stock Exchange Composite is a modified market capitalization weighted index comprising all of the stocks listed on the Jakarta Stock Exchange.



Note. The horizontal lines in each graph represent (Q1-3*IQR, Q3+3*IQR), where Q1 is the first quartile, Q3 is the third quartile and IQR=Q3-Q1.

Fig. 1. Time plots of log return.

box plot.[10] All markets, except for the Taiwanese, show extreme outliers around the time of Asian crisis in 1997. It is also found that many markets exhibit a substantially higher degree of volatility after the crisis.[11] Outliers occurred in the Malaysian market in the second half of 1998, which are related to political unrest and imposition of capital control. Outliers occurred in the Philippine market in late 2000 and early 2001 are also related to political events.[12]

For weekly data, we have considered the whole sample as well as two sub-samples. The sub-sample I is from 1990 to the end of 1996 (364 observations) and sub-sample II from January 1998 to 2005 (381 observations). This is to separate the effect of the Asian crisis, which influenced the stock markets of many Asian countries. The holding periods (k_i's) considered for weekly data are (2, 4, 8, 16, 32) for all cases. For daily data, the test is conducted with moving sub-sample windows of 1000 observations, with the holding periods set to (2, 5, 10, 20, 40). We use relatively short holding periods because variance ratio tests are mainly used over a short horizon (see, for example, Ayadi and Pyun, 1994; Chang and Ting, 2000). However, we also have conducted the tests with much longer holding periods, namely holding periods of up to 90 days for daily data and up to 80 weeks for weekly data. The results are not reported as they are found to be qualitatively similar to those reported in this paper, especially for the case of daily data. To reduce the burden of computation, the sub-sample is taken at every Wednesday, starting from the 3rd November, 1993. For the wild bootstrap test (MV*), the number of bootstrap replications m is set to 2000.

[10] See, for example, De Veaux, Velleman, and Bock (2007), chapter 5.

[11] The Chow test has been conducted for the shift of volatility with the break point of July 1997 using daily returns. We have found that all markets show positive shift of volatility after the crisis, except for Japanese and Philippine markets that show no shift and for Taiwanese market that show a negative shift. We have also conducted CUSUM-type tests of Andreou and Ghysels (2002) that provided similar results. We would like to thank Bill Zhang for sharing his GAUSS codes for the computation of these CUSUM-type tests.

[12] Further details of these events can be found at Campbell Harvey's website at Duke University, which provides chronological lists of important events for emerging markets (http://www.duke.edu/~charvey/Country_risk/couindex.htm).

5. Empirical results

Table 4 reports the multiple VR test results weekly data. The 5% and 10% critical values for the Chow–Denning (MV) and joint sign (JS) tests are given at the bottom of Table 4, while the *p*-values are reported for the wild bootstrap (MV*) test. Fig. 2a to c report the results of the MV, MV* and JS tests for daily data with the sub-sample window of 1000 observations. Similarly to the case of weekly data, the test statistics are plotted for the MV and JS tests, while *p*-values for the MV* test.

Table 4
Multiple VR test results for weekly data

	MV	JS	MV*
	(test statistics)		(*p*-values)
HK			
Sub-sample I	0.86	1.74	0.76
Sub-sample II	1.53	0.94	0.31
Whole Sample	2.01	2.01	0.12
JAP			
Sub-sample I	0.95	1.01	0.71
Sub-sample II	1.01	0.93	0.67
Whole Sample	0.47	1.12	0.97
KOR			
Sub-sample I	0.68	2.20*	0.89
Sub-sample II	1.11	1.28	0.62
Whole Sample	1.10	2.79**	0.60
TAW			
Sub-sample I	1.32	0.61	0.41
Sub-sample II	0.82	1.29	0.78
Whole Sample	1.58	1.08	0.28
IND			
Sub-sample I	6.49**	5.80**	0.00**
Sub-sample II	3.28**	1.48	0.00**
Whole Sample	5.80**	5.43**	0.00**
MAL			
Sub-sample I	0.75	1.66	0.86
Sub-sample II	2.10	2.19*	0.07*
Whole Sample	1.26	2.26*	0.45
PHI			
Sub-sample I	2.30	1.61	0.05*
Sub-sample II	2.95**	1.75	0.01**
Whole Sample	3.05**	2.31*	0.01**
SIN			
Sub-sample I	0.83	0.74	0.83
Sub-sample II	0.95	3.15**	0.77
Whole Sample	0.74	2.02	0.81
THI			
Sub-sample I	1.26	2.52**	0.48
Sub-sample II	2.77**	1.45	0.01**
Whole Sample	3.01**	2.37*	0.01**

MV: Chow–Denning test, MV*: wild bootstrap test, JS: joint sign test; The 5% and 10% critical values for the MV test are 2.56 and 2.31 respectively. The 5% and 10% critical values for the JS test are 2.39 and 2.11 respectively when sample size is 800; and 2.32 and 2.04 when the sample size is 375. The symbol “*” indicates the significance at 10% level of significance while “**” the significance at 5%.

First, we consider the group that includes Hong Kong, Japanese, Korean, and Taiwanese markets (Group 1). When weekly data is used, the markets of Hong Kong, Japan, and Taiwan do not show any rejection of the martingale hypothesis before and after the Asian crisis, as well as for the whole sample period. The Korean market shows rejection before the crisis, only when the JS test is used. When the MV* test is used for the daily data (Fig. 2a), the martingale cannot be rejected for all markets in this group. The Japanese and Korean markets show departure from the martingale around 1999, but this seems to be an effect of influential outliers occurred as a result of the Asian crisis. Similar results are evident when the MV test is used. If the JS test is used, the Hong Kong, Korean and Japanese markets show the

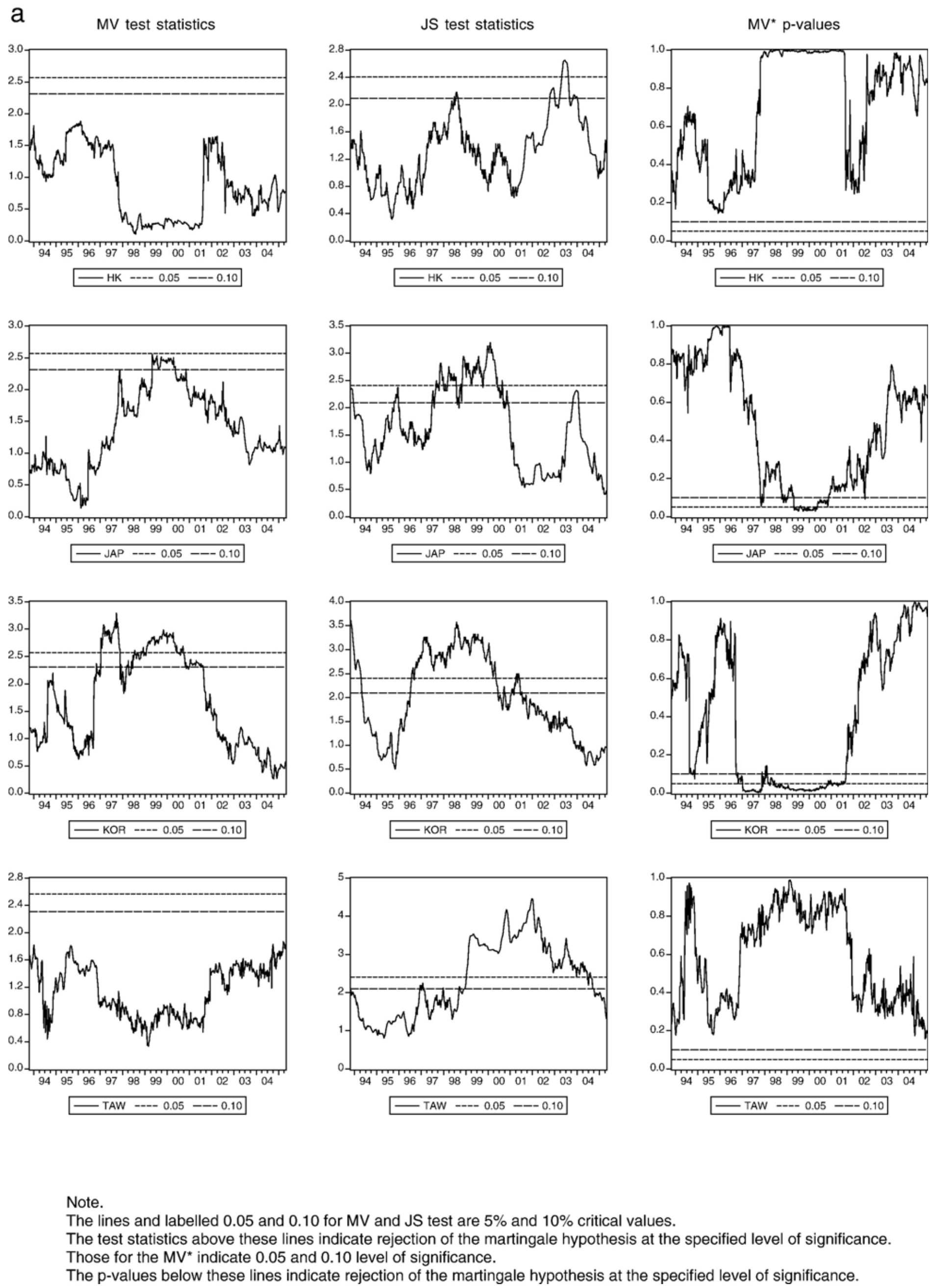


Fig. 2. a. Multiple variance ratio tests for daily data (Group 1 markets). b. Multiple variance ratio tests for daily data (Group 2 markets). c. Multiple variance ratio tests for daily data (Group 3 markets).

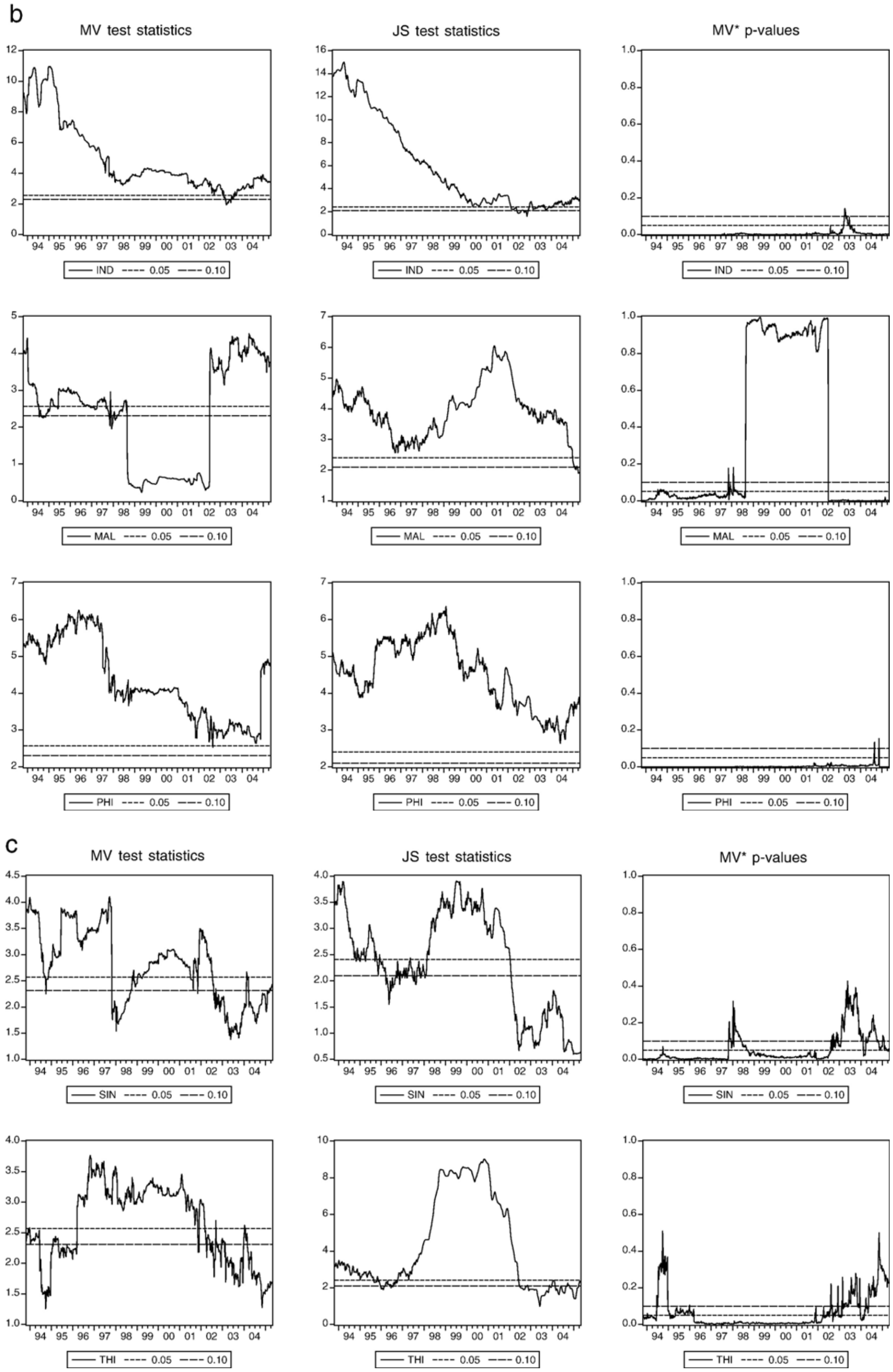


Fig. 2 (*continued*).

martingale behavior overall. With the JS test, the Taiwanese market shows rejection of the martingale hypothesis after the Asian crisis, but it shows the martingale behavior after 2003. Hence, based on these results, we conclude that the markets in the first group are efficient in the weak-form, both before and after the Asian crisis.

Next, we consider the group consists of Indonesian, Malaysian, and Philippine markets (Group 2). From Table 4, the efficiency of the Indonesian market is soundly rejected by all three tests when weekly data is used, both before and after the crisis. The same conclusion is evident when daily data is used, as the martingale hypothesis is rejected for nearly the entire period for Indonesian market in Fig. 2b. The efficiency of Malaysian market is not rejected when weekly data is used before the crisis, but it is rejected by the JS and MV* test after the crisis. If daily data is used, the martingale hypothesis for the Malaysian market is clearly rejected by all three tests, both before and after the crisis. For Malaysian market, the MV and MV* tests show abrupt changes in the inferential outcomes in 1998 and 2002. This is due to the influential outliers occurred in 1998. The JS test does not show such abrupt changes because the sign test is completely free from the effect of outliers. Using weekly data, the efficiency of the Philippine market is rejected by the MV* test before the crisis, and by the MV and MV* tests after the crisis. If daily data is used, all three tests soundly reject the martingale hypothesis for the Philippine market. It appears that the outliers occurred in the Philippine market around 2001 were not influential to the outcomes of the test, in contrast to the case of Malaysian market. Overall, we conclude that these three markets have been inefficient since early nineties.

The last group (Group 3) includes Singaporean and Thai markets. Using weekly data, the efficiency of the Singaporean market is rejected by the JS test after the crisis. But the examination of the results from daily data from Fig. 2c shows the reverse result: all three tests show rejection before the crisis, but overall acceptance of the martingale hypothesis afterwards. The Thai market shows similar results. From weekly data, the efficiency of the Thai market is overall rejected, both before and after the crisis. But according to the daily data, all three tests indicate the rejection before the crisis but acceptance of the martingale hypothesis afterwards. Table 5 presents a summary of overall inferential outcomes for both weekly and daily data before and after the crisis.

The weekly and daily data results are largely consistent for the first and second groups of markets. For Singaporean and Thai markets, however, the results for the post-crisis period are somewhat unexpected as one would expect evidence of efficiency from low frequency data (weekly) and inefficiency from high frequency data (daily). However, it is likely that the results from daily data are the outcomes of more powerful testing, as they are based on a larger sample size (see Section 3.4). Moreover, the use of moving sub-sample window for daily data is likely to have yielded results more robust to structural change and outliers than those from weekly data. This suggests that the Singaporean and Thai markets have become efficient have become efficient after the crisis.

It is interesting to compare the results obtained in this study with those from past studies. For the first group, past studies in general found the Hong Kong, Korean, Taiwanese markets to be inefficient, while this study provides evidence in favor of efficiency of these markets. For example, we have found the Taiwanese market has been efficient, in contrast with the findings of Chang and Ting (2000) and Peng et al. (2004). The findings for the second group of the markets (Indonesia, Malaysia and Philippine) from the current and past studies are in agreement. As for the last group, past studies found that the Singaporean and Thai markets to be inefficient before the Asian crisis, as with the findings of this study (see, for example, Lima and Tabak, 2004). However, we have found new evidence that these two markets have become efficient after the crisis. This result is also different from the findings of Hoque et al. (in press) that the Asian crisis has exerted little influence on the nature of market efficiency of these two markets. We suspect that our

Table 5
Overall inferential outcomes of the martingale hypothesis

	Weekly		Daily	
	Before crisis	After crisis	Before crisis	After crisis
HOK	A	A	A	A
JAP	A	A	A	A
KOR	R	A	A	A
TAW	A	A	A	A
IND	R	R	R	R
MAL	A	R	R	R
PHI	R	R	R	R
SIN	A	R	R	A
THI	R	R	R	A

A: Acceptance, R: Rejection.

results are different from those of past studies, due to a number of reasons including the use of more disaggregated data, more powerful tests which are robust to outliers and structural changes, and updated data.

According to Kawakatsu and Morey (1999), many of the emerging markets considered in this study implemented financial liberalization measures since late eighties or early nineties. Although Kawakatsu and Morey (1999) found little evidence that the financial liberalization improves market efficiency, they noted that liberalization is a gradual process and may take some time to show its effect. Our analysis using updated data also has found little evidence that financial liberalization has improved market efficiency of Indonesian, Malaysian, and Philippine markets. However, we have found evidence that financial liberalizations that followed the Asian crisis improved the market efficiency of the Singaporean and Thai markets.

This paper provides strong evidence for stock market efficiency in the developed equity markets of Hong Kong and Japan and an advanced emerging market of Taiwan.[13] Although South Korea is categorized as an advanced emerging market, in general, this market was found to be efficient. For Singapore — a developed market, this paper found evidence of market efficiency after the crisis. The secondary emerging equity markets of Indonesia, Malaysia, Philippines and Thailand were subject to pricing inefficiency. Among these markets, we find that only Thai market has become efficient after the Asian crisis. In general, the regulatory framework and business culture in these countries were not conducive of good and transparent corporate governance, which adversely affected the quality and speed of transmission of price-sensitive information in these markets (see OECD, 2003). Thus, despite continuing financial liberalization in Indonesia, Malaysia and Philippines, the degree of informational efficiency appears to remain small in these markets. This finding is compatible with an earlier study on emerging financial markets by Kawakatsu and Morey (1999). Thus, the results indicate that market efficiency may depend on the level of equity market development as well as the regulatory framework conducive of transparent corporate governance.

6. Conclusion

This paper examined whether a group of selected Asian stock market returns (Hong Kong, Indonesia, Japan, Korea, Malaysia, Philippines, Taiwan, Thailand and Singapore) follow a martingale process. This exercise is important because the martingale property has strong implications to the stock market efficiency in the weak-form. In this study, new multiple variance ratio tests based on the wild bootstrap and signs are adopted, in addition to the conventional Chow–Denning test. According to our Monte Carlo results, these new variance ratio tests have superior small sample properties to those of the Chow–Denning test. We have used daily and weekly price indices from 1990 to 2005 in this paper.

The paper finds that market efficiency varies with the level of equity market development. In general, the developed or advanced emerging markets (Hong Kong, Japan, Korea, Singapore, Taiwan) show weak-form efficiency, while the secondary emerging markets (Indonesia, Malaysia, Philippines) are found to be inefficient. We have found evidence that Singaporean and Thai markets have become efficient after the Asian crisis in 1997. Despite financial market liberalization, the secondary emerging markets have shown little sign of market efficiency. Further research can be conducted to identify the economic and institutional factors underlying our empirical findings and to determine whether or not any pricing inefficiency can be exploited economically to earn a risk-adjusted abnormal return.

References

Andreou, E., Ghysels, E., 2002. Detecting multiple breaks in financial market volatility dynamics. Journal of Applied Econometrics 17, 579–600.

Ayadi, O.F., Pyun, C.S., 1994. An application of variance ratio test to the Korean securities market. Journal of Banking and Finance 18, 643–658.

Belaire-Franch, G., Contreras, D., 2004. Ranks and signs-based multiple variance ratio tests. Working paper. University of Valencia. http://les1.man.ac.uk/sapcourses/esgc/Papers2004/Belaire-Franch.pdf.

Belaire-Franch, G., Opong, K.K., 2005a. A variance ratio test of the behaviour of some FTSE equity indices using ranks and signs. Review of Quantitative Finance and Accounting 24, 93–107.

Belaire-Franch, G., Opong, K.K., 2005b. Some evidence of random walk behavior of Euro exchange rates using ranks and signs. Journal of Banking and Finance 29, 1631–1643.

Berkowitz, J., Killian, L., 2000. Recent developments in bootstrapping time series. Econometric Reviews 19, 1–48.

Berneburg, M., 2004. Are European equity style indexes mean reverting? Testing the validity of the efficient market hypothesis. Discussion Paper, vol. 193. Institut Fûr Wirtschaftsforschung Halle, Halle, Germany.

Black, F., 1986. Noise. Journal of Finance 41, 529–543.

[13] As noted at the end of Section 1, this study uses the FTSE country classification to determine the level of equity market development.

Chang, K.-P., Ting, K.S., 2000. A variance ratio test of the random walk hypotheses for Taiwan's stock market. Applied Financial Economics 10, 525–532.

Chow, K.V., Denning, K.C., 1993. A simple multiple variance ratio test. Journal of Econometrics 58, 385–401.

Coggin, T.D., 1998. Long memory in equity style indexes. The Journal of Portfolio Management 37–46 (Winter).

De Veaux, R., Velleman, P., Bock, D., 2007. Stats: Data and Models: International Edition, 2e. Pearson/Addison Wesley, Boston.

Dedondt, W., Thaler, R., 1985. Does the stock market overreact? Journal of Finance 40, 793–808.

Efron, B., 1979. Bootstrap methods: another look at the jackknife. Annals of Statistics 7, 1–26.

Fama, E., 1965. The behaviour of stock market prices. Journal of Business 38, 34–105.

Fama, E., 1970. Efficient capital markets: a review of theory and empirical work. Journal of Finance 25, 383–417.

Fama, E., French, K., 1988. Permanent and temporary components of stock prices. Journal of Political Economy 47, 246–273.

Fong, W.M., Koh, S.K., Ouliaris, S., 1997. Joint variance–ratio tests of the martingale hypothesis for exchange rates. Journal of Business and Economic Statistics 15, 51–59.

Hoque, H., Kim, J.H., Pyun, C.S., in press. A comparison of variance ratio tests of random walk: a case of Asian emerging stock markets. International Review of Economics and Finance.

Huang, B.-N., 1995. Do Asian stock market prices follow random walks? Evidence from variance ratio test. Applied Financial Economics 5, 251–256.

Kawakatsu, H., Morey, M.O., 1999. An empirical examination of financial liberalization and the efficiency of emerging market stock prices. Journal of Financial Research XXII, 385–411.

Kendall, M., 1953. The analysis of economic time series. Journal of the Royal Statistical Society Series A 96, 11–25.

Kim, J.H., 2006. Wild bootstrapping variance ratio tests. Economics Letters 92, 38–43.

Lee, C., Chen, G., Rui, O., 2001. Stock returns and volatility on China's stock markets. Journal of Financial Research 24, 523–543.

Lehmann, B.M., 1990. Fads, martingales and market efficiency. Quarterly Journal of Economics 105, 1–28.

Li, H., Maddala, G.S., 1996. Bootstrapping time series models. Econometric Reviews 15, 115–158.

Lima, E.J.A., Tabak, B.B., 2004. Tests of the random walk hypothesis for equity markets: evidence from China, Hong Kong and Singapore. Applied Economics Letters 11, 255–258.

Lo, A.W., 1991. Long-term memory in stock market prices. Econometrica 59, 1279–1313.

Lo, A.W., MacKinlay, A.C., 1988. Stock market prices do not follow random walks: evidence from a simple specification test. The Review of Financial Studies 1, 41–66.

Lo, A.W., MacKinlay, A.C., 1989. The size and power of the variance ratio test in finite samples. Journal of Econometrics 40, 203–238.

Mammen, E., 1993. Bootstrap and wild bootstrap for high dimensional linear models. The Annals of Statistics 21, 255–285.

Mun, F.W., Kee, K.S., 1994. Do Asian stock market prices follow martingales? Evidence from spectral shape tests. Asia Pacific Journal of Management 11, 345–359.

OECD, 2003. White paper on corporate governance in Asia. OECD Publications. Paris.

Peng, K.-L., WU, C.-H., Goo, Y.-J., 2004. The development of a new statistical technique for relating financial information to stock market returns. International Journal of Management 21, 492–505.

Poterba, J.M., Summers, L.H., 1988. Mean reversion in stock returns: evidence and implications. Journal of Financial Economics 22, 27–59.

Richardson, M., 1993. Temporary components of stock prices: a skeptic's view. Journal of Business & Economic Statistics 11, 199–207.

Richardson, M., Smith, T., 1991. Tests of financial models in the presence of overlapping observations. The Review Financial Studies 4, 227–254.

Ryoo, H.J., Smith, G., 2002. Korean stock prices under price limits: variance ratio tests of random walks. Applied Financial Economics 12, 475–484.

Stangle, B., 2005. Market efficiency versus behavioral finance. Journal of Applied Corporate Finance 17, 124–134.

Savin, N.E., 1984. Multiple hypothesis testing. In: Griliches, Z., Intriligator, M.D. (Eds.), Handbook of Econometrics. vol. 2. North-Hollan, Oxford. Chapter 14.

Stoline, M.R., Ury, H.R., 1979. Tables of the Studentized maximum modulus distribution and an application to multiple comparisons among means. Technometrics 21, 87–93.

Whang, Y.-J., Kim, J., 2003. A multiple variance ratio test using subsampling. Economics Letters 79, 225–230.

White, H., 1999. Asymptotic Theory for Econometricians. Academic Press, New York.

Wright, J.H., 2000. Alternative variance–ratio tests using ranks and signs. Journal of Business & Economic Statistics 18, 1–9.

Yilmaz, K., 2003. Martingale property of exchange rates and central bank intervention. Journal of Business & Economic Statistics 21, 383–395.