Review of Pacific Basin Financial Markets and Policies
Vol. 6, No. 2 (2003) 199–222

Study on the Non-Random and Chaotic Behavior of Chinese Equities Market

Patrick K. K. Chu
Faculty of Business Administration, University of Macau, Macao SAR, China
PatrickC@umac.mo

After the stock market crash of October 19, 1987, interest in nonlinear dynamics and chaotic dynamics have increased in the field of financial analysis. The extent that the daily return data from the Shanghai Stock Exchange Index and the Shenzhen Stock Exchange Index exhibit non-random, nonlinear and chaotic characteristics are investigated by employing various tests from chaos theory. The Hurst exponent in R/S analysis rejects the hypothesis that the index return series are random, independent and identically distributed. The BDS test provides evidence for nonlinearity. The estimated correlation dimensions provide evidence for deterministic chaotic behaviors.

Keywords: Chaos theory; rescaled range analysis; Hurst exponent; BDS test; correlation dimension estimation; Shenzhen Stock Exchange; Shanghai Stock Exchange.

1. Introduction

The two stock markets in Mainland China, the Shanghai Stock Exchange (SHSE) and the Shenzhen Stock Exchange (SZSE) began their operations in December 1990 and July 1991, respectively. In these years, both stock markets have dramatically expanded. In the Shanghai Stock Exchange, by the end of 2002, there were 728 companies listed with a trading volume of 2,825 billion yuan. In the Shenzhen Stock Exchange there were 508 companies listed, with a trading volume of 1,296.5 billion yuan by the end of 2002. Both the Shanghai Stock Exchange and Shenzhen Stock Exchange were closed markets at the early stage of operation. In order to attract foreign capital, both stock markets developed the Class B Shares Exchange for the non-Mainland Chinese investors in 1992. The Class B shares are traded by U.S. dollar in the Shanghai Stock Exchange and by HK dollar in the Shenzhen Stock Exchange, respectively. Meanwhile, the trading of Class A shares are restricted to Mainland Chinese investors. After May 2001, the

Mainland Chinese are also allowed to trade Class B shares with their authorized foreign currencies.

One of the challenges in the capital market theory today is to develop theories that are capable of explaining the movements in asset prices and returns. According to Fama's (1970) efficient market hypothesis (EMH), all relevant public (or private) information has been reflected in the current prices. Many empirical studies in the 1970s and 1980s seemed to support the EMH. However, most of the recent studies show that the EMH fails. Before chaos theory has been applied in the field of economics or finance, the approaches to describing the asset price behavior were by random walk model with uncorrelated innovations (Fama, 1970). Other deviations from the random walk would then be explained by the anomalies, such as weekend effect, January effect and neglected-firm effect, etc. Recent research into an explanation of stock return behavior has drawn on the field of nonlinear dynamics, including chaos theory.

Chaos theory is based on the assumptions that the underlying system is a *nonlinear* and *deterministic* process. Some findings have pointed out that linear models are not very good in trying to capture the complexities of the economic system. In fact, a number of recent studies have found strong evidence of nonlinearity in the short-term movements of asset returns (Hsieh, 1993). During these years, nonlinear analysis, originally developed in physics and the natural sciences, is rapidly expanding in different research areas. Finance and economics are areas that strongly need the application of such approach because empirical researches show that the linear models are not adequate to explain the underlying dynamics of recent economic data. The chaotic and nonlinear deterministic systems can attract the attention of many scholars from different fields because of two reasons. First, it is relatively easy to tune the parameters of dynamic economic models so that they generate complex dynamics (Benhabib and Nishimura, 1979). Second, it is easy to construct examples of nonlinear dynamics that appear random in linear tests (Sakai and Tokumaru, 1980).

The existence of chaotic behaviors can be detected by the following indications:

1. With the chaotic behavior, there will be an existence of strange attractors characterized by fractal shape.
2. Chaotic process is not random but independent and identically distributed (i.i.d.).

3. Chaotic process is a *nonlinear* process.
4. Chaotic process is a *deterministic* process which will retain its dimensionality when it is placed in a higher embedding dimension.
5. Chaotic process is sensitive to initial conditions.

Recently, there are extensive researches done on chaotic behavior and nonlinear dynamics of stock price behavior in industrial nations such as U.S. and U.K. However, there is a lack of similar studies for the Chinese stock markets. As the EMH and the random walk model fail to capture the behaviors in many stock markets, and some researchers like [Hsieh (1991, 1993) and Brock *et al.* (1991)] have pointed out that nonlinear dynamics is appropriate to explain the complexity in many stock returns series. Fama (1965) admits that linear modeling techniques are limited in capturing the complicated patterns which chartists claim to see in stock prices.

The Chinese Stock Markets have been found to be non-efficient (Huo, 1996). Papaioannou and Karytinos (1995) indicated that a non-efficient market is very suitable for chaotic analysis, because the number of underlying moving forces (degree of freedoms in statistical terms) is fewer than those in developed and more efficient markets such as New York Securities Exchange and London Securities Exchange.

At present, there is no single reliable statistical test for chaos. A feasible way to detect nonlinear and chaotic dynamics in time series is to adopt various tests available, in order to avoid misleading results and conclusions (Papaioannou and Karytinos, 1995).

The objective of this paper is to examine whether the time series of the Composite Index returns in Shanghai and Shenzhen Stock Exchange are generated by nonlinear dynamical system. If so, the predictive ability of the system is strongly limited, especially for long-run predictions. Furthermore, we focus on examining whether chaos theory can capture the complex and random behaviors that may not be obtained or derived from a stochastic approach or the traditional random walk model.

The rest of this paper is organized as follows: Section 2 provides a review of previous studies on non-random and chaotic behaviors in economic and financial series. Section 3 illustrates the formulae, interpretations and usages of the methodologies that will be employed in this study. Section 4 describes the data that have been examined in this study. Section 5 presents the empirical results. Section 6 is devoted to general conclusions and implications of the financial analysts.

2. Previous Studies

Some studies to test the evidence of identical and independent distributed characteristics in equity returns have been done in recent years. Hsieh (1991) has doubt on the behavior of returns on the S&P 500 Index following a random walk. Peters (1992) shows that the S&P 500 Index monthly return series has long-term dependence and a cycle length of 48 months. Peters (1994) finds that the same phenomenon appears in Dow Jones Industrial Index. Hampton (1996) report a strong and continuing dependence in the S&P 500 returns series before the market crashes in 1987 and 1990. Opong *et al.* (1999) find the same evidence of non-i.i.d. in London Financial Times Stock Exchange (FTSE) returns series. Same results are also reported in the other European and some emerging markets, Errunza *et. al.* (1994) find evidence of fractal dynamics in equity returns of Germany, Japan, Argentina, Brazil, Chile, India and Mexico. Papaioannou and Karytinos (1995) find that the Athens Stock Market Index returns series have dependence. Not only stock market returns, but also security returns have been found to have evidence of dependence. Greene and Fieltz (1977) report that 200 security return series listed on the New York Stock Exchange are characterized by long-term dependence.

With the massive computer power available, there is no need to make any simplifying assumptions and this results in the evolution of more complicated models. The mathematical models rapidly become complicated and nonlinear. The fact that most financial series are nonlinear dependent has been found. In U.S. equity market returns, Hsieh (1991) finds that the S&P 500 weekly and daily return series are nonlinearly dependent. Lin (1997) also report same evidence of nonlinear dependence in the Dow Jones Industrial Average return series. The dependence is more significant in 1995 when the U.S. equity markets made the start of a strong bull run. Pandey *et al.* (1997) report evidence of nonlinear dependence in the index returns of Hong Kong, Japan and the U.S. Regarding the emerging markets, Hamill and Opong (1997) report nonlinear dependence in Irish stock returns, while Papaionnou and Karytinos (1995) find that the Athens Stock Exchange Index returns are also nonlinearly dependent.

One of the indications of a chaotic process is that the process is deterministic. Arguments on whether deterministic structure appears in the economic and financial series are being continued in recent years. Up to now, most empirical findings are contradictory. Scheinkman and LeBaron (1989) report the evidence of determinism in weekly stock returns. Willey

(1992) finds the deterministic characteristic appears in both daily S&P 100 Index returns and NASDAQ Index returns. On the other hand, Howe *et al.* (1997) find no evidence of deterministic patterns in Australia and Hong Kong equity returns.

3. Research Methodologies

Currently, there is no single reliable statistical test for the existence of chaotic behavior. Combining various tests is a common practice in some other similar studies. The indications of chaotic behavior can be detected by the following methodologies, which are well developed in the fields of physics and the natural sciences. In this study, they will be applied to finance and will be used to detect the existence of chaotic behavior in the time series of Chinese stock market indices.

3.1. *Rescaled range (R/S) analysis*

The efficient market hypothesis (EMH) assumes that all investors immediately react to new information, so that the future price movement of a stock is unrelated to the past or present patterns of price movements. Actually, do people behave in this manner? Peters (1991a) mentioned that most people wait for confirming information and do not react until a trend is clearly established. Consequently, there will be an uneven assimilation of information. This will cause the stock price movement to follow a biased random walk, instead of random walk. Non-random walk implies that there is memory underlying in the series. Whether the stock price movement follows a random walk or not can be detected by the *rescaled range analysis* or R/S analysis. The R/S analysis is an ideal statistical tool for analyzing the occurrence of rare events and is robust to possible nonlinear process that normality assumption may not be needed. Due to this, the R/S analysis should be chosen to describe the stock market crashes. The result of the R/S analysis is the *Hurst exponent*, which is a measure of the bias or trend in a time series.

The R/S analysis was developed by hydrologist H. E. Hurst in 1951. His work was derived from Einstein's work regarding the Brownian motion of physical particles to deal with the problem of reservoir control near Nile River Dam. Peters (1994) has applied the R/S analysis to analyze both periodic and nonperiodic cycles, following the procedures for calculating Hurst exponent used by Peters (1994), beginning with the smallest value of n, the partition size, to partition the series into A sequential non-overlapping blocks, $A = \frac{N}{n}$, where A is the number of partition, N is the amount of data in the series and n is the amount of data in each partition. The data in each block is $x_{t,a}$. The mean of $x_{t,a}$ ($\overline{x_a}$) for the ath block of data is defined as

$$\overline{x_a} = \left(\frac{1}{n}\right)\sum_{t=1}^{n} x_{t,a}\,. \tag{1}$$

The accumulated differences, X_a, between each $x_{t,a}$ and the mean for each block of data, $\overline{x_a}$, given by

$$X_a = \sum_{t=1}^{n}(x_{t,a} - \overline{x_a})\,. \tag{2}$$

The range, R_a, is defined as the difference between the minimum and the maximum cumulative deviation for each block of data, given by

$$R_a = \max(X_a) - \min(X_a)\,. \tag{3}$$

The standard deviation, S_a, for each block of data is determined, given by

$$S_a = \sqrt{\frac{\sum_{t=1}^{n}(x_{t,a} - \overline{x_a})^2}{n}}\,. \tag{4}$$

The average rescaled range $(R/S)_n$ for length n, which is computed for different lengths until $n = N/2$, is defined as

$$(R/S)_n = \left(\frac{1}{A}\right) \times \sum_{a=1}^{A}\left(\frac{R_a}{S_a}\right). \tag{5}$$

The final step is to apply an OLS regression with $\ln(R/S)_n$ as the dependent variable and $\ln(n)$ as the independent variable through the plot. The regression coefficient of the regression equation (6), provides the estimation of H, the Hurst exponent

$$\ln(R/S)_n = H \cdot \ln(n) + C\,. \tag{6}$$

The value of H can be interpreted in the following ways:

- $H = 0.50$: denotes a random and statistically independent (uncorrelated) series — a random walk. The present does not influence the future. The correlation coefficient is 0. Its probability density function is normal. Such process increases with the square root of time.
- $0.50 < H \leq 1.00$: denotes a "persistent", or trend-reinforcing series. That is, the data contains long-term memory and has a tendency to follow the current trend in the next period. This process is said to be mean-averting.

For example, if the series has been up (down) in the last period, then it is more likely to be up (down) in the next period. The strength of the persistence increases and the correlation coefficient approaches 1 as H approaches 1. Such persistent series are said to be fractional Brownian motion (biased random walks), in terms of nonlinear dynamics, the series displays sensitive to initial conditions. Peters (1989) mentioned that the higher H is the stronger the persistence and the less "white noise" in the time series.

- $0 \leq H < 0.50$: denotes an "anti-persistent", or ergodic series. That is, the data has a tendency to reverse the current trend. This process is said to be mean-reverting. For example, if the system has been up (down) in the previous period, it is more likely to be down (up) in the next period. The strength of the anti-persistence depends on how close H is to zero and the correlation coefficient approaches -1 as H approaches 0. The frequent reversal results in less distance covered by the process than would occur given a random process.

An additional statistic used is the V-statistic. The basis of V-statistic is that a periodic system corresponds to a limit cycle or a similar type of attractor, such that its phase space portrait would be a bonded set. Because the range could never grow beyond the amplitude, the R/S values would reach a maximum value after one cycle. The V-statistic is given by

$$V_n = \frac{(R/S)_n}{\sqrt{n}}\,. \tag{7}$$

In the V_n versus $\log(n)$ plot, it will be flat if the process is an independent, random process ($H = 0.5$). If the process is persistent ($H > 0.5$), the graph will be upwardly sloping. Conversely, if the process is anti-persistent ($H < 0.5$), the graph will be downward sloping. The cycle length, can be discerned from the "break-point" in this plot to occur when V reaches a peak and then flattens out.

3.2. *Phase space reconstruction and embedding dimension*

The phase space reconstruction is the basis of the other methodologies such as BDS test and correlation dimension estimation.

Suppose that information is available on a univariate time series and it is known that this series is part of a larger, n-dimensional deterministic economic model, i.e. there are n variables. Martin *et al.* (1994) indicated that when testing for the presence of non-linearity and hence need to identify the dynamics of this system, it is not necessary to have time series data on the remaining $n-1$ variables in the system if the available time series is embedded. Besides, Takens (1981) has proved that an embedded univariate time series can encapsulate the information of a multivariate time series model without any loss of information. Takens (1981) suggested that the phase space reconstruction for a series of N observations $\{x_i\} = [x_1, x_2, \ldots, x_{N-1}, x_N]$ can be done by transforming this series of observations into a series of scalar vectors, and is given by

$$\begin{aligned} x_1^m &= (x_1, x_2, \ldots, x_m) \\ x_2^m &= (x_2, x_3, \ldots, x_{m+1}) \\ &\vdots \\ x_{N-m}^m &= (x_{N-m}, x_{N-m+1}, \ldots, x_N) \end{aligned} \tag{8}$$

where the parameter m is the "embedding dimension".

3.3. *BDS test*

BDS test was suggested by Brock, Dechert and Scheinkman (1987). This hypothesis testing uses the test statistics in which mechanism is based on the correlation integrals. The BDS test is a powerful tool for detecting serial dependence in time series. It tests the null hypothesis of independent and identically distributed (I.I.D.) against an unspecified alternative. The null and alternative hypothesis are as follows:

H_0: The data are independently and identically distributed (I.I.D.).
H_1: The data are not I.I.D.; this implies that there may be some serial dependence. If the linear dependence has been removed in the time series, the serial dependence is thus nonlinear.

However, BDS test is unable to distinguish between nonlinear deterministic chaos and nonlinear stochastic systems. BDS test cannot test chaos directly but only nonlinearity, provided that any linear dependence has been removed from the data (e.g. using traditional ARIMA-type models or taking a first difference of natural logarithms). Nevertheless, since a *nonlinear* process is one of the indications of chaos, we may use the BDS test to detect such indication.

Given an embedded series with N (length of original time series) — m (embedding dimension) observations $\{x_i^m\} = [x_1^m, x_2^m, x_3^m, \ldots x_{N-m}^m]$, the correlation integral that measures the spatial correlation among the points can be computed by adding the number of pairs of points (i, j), which are "close" in the sense that the points are within a radius or tolerance ε of each other:

$$C_{\varepsilon,m} = \frac{1}{N_m(N_m - 1)} \sum_{i \neq j} I_{i,j;\varepsilon} \,. \tag{9}$$

The above equation may be read as: the correlation integral for a time series of length N that is embedded in m-dimensional space with a correlation distance of ε. The indicator function $I_{i,j;\varepsilon}$ in the above equation is calculated by the following equation:

$$I_{i,j;\varepsilon} = \begin{cases} 1 & \text{if } \|x_i^m - x_j^m\| \leq \varepsilon \\ 0 & \text{otherwise}\,. \end{cases} \tag{10}$$

$C_{\varepsilon,m}$ measures the probability that any particular pairs in the time series are close. Brock, Dechert, and Scheinkman (1987) showed that if the time series $\{x_t\}$ is I.I.D., then:

$$C_{\varepsilon,m} \approx [C_{\varepsilon,1}]^m \,. \tag{11}$$

Simply stated, if the series is I.I.D., the correlation integral at an embedding dimension (m) and given a certain value of ε, can be approximated by the mth power of $C_{\varepsilon,1}$, where the correlation integral is under the embedding dimension equal to 1.

If the requirements of having ratios $\frac{N}{m}$ greater than 200, allowing the values of $\frac{\varepsilon}{\sigma}$ to range from 0.5 to 2 and the values of m from two to five are fulfilled, the quantity $[C_{\varepsilon,m} - (C_{\varepsilon,1})^m]$ has an asymptotic normal distribution with a zero mean and a variance $V_{\varepsilon,m}$ defined as:

$$V_{\varepsilon,m} = 4[K^m + 2\sum_{j=1}^{m-1} K^{m-j} C_\varepsilon^{2j} + (m-1)^2 C_\varepsilon^{2m} - m^2 K C_\varepsilon^{2m-2}] \tag{12}$$

where

$$K = K_\varepsilon = \frac{6}{N_m(N_m - 1)(N_m - 2)} \sum_{i<j<N} h_{i,j,N;\varepsilon} \,, \quad \text{and}$$

$$h_{i,j,N;\varepsilon} = \frac{[I_{i,j;\varepsilon} I_{j,N;\varepsilon} + I_{i,N;\varepsilon} I_{N,j;\varepsilon} + I_{j,i;\varepsilon} I_{i,N;\varepsilon}]}{3}$$

$$I_{i,j;\varepsilon} = \begin{cases} 1 & \text{if } \|x_i^m - x_j^m\| \leq \varepsilon \\ 0 & \text{otherwise}\,. \end{cases}$$

The BDS test statistic is defined as

$$BDS_{\varepsilon,m} = \frac{\sqrt{N}[C_{\varepsilon,m} - (C_{\varepsilon,1})^m]}{\sqrt{V_{\varepsilon,m}}} \,. \tag{13}$$

This test will be repeated at different values of ε and m. Lin (1997) suggested that the appropriate values of $\frac{\varepsilon}{\sigma}$ range from 0.5 to 2, Brock *et al.* (1987) pointed out the appropriate values of m are between 2 and 5.

Since BDS test is a two-tailed test, we should reject the null hypothesis if the BDS test statistic is greater than the positive critical z-value or less than the negative critical z-value. For example, if $\alpha = 0.05$, the critical z-value $= \pm 1.96$.

If the alternative hypothesis of dependence is accepted, then testing between deterministic chaos and stochastic processes can proceed by the estimation of correlation dimension, which will be discussed in the next section.

3.4. *Correlation dimension estimation*

The correlation dimension estimation, the most common method used to determine the fractal dimension of a system, is used to differentiate between deterministic chaos and stochastic systems. As mentioned in Sec. 2, one of the indications of chaos is that the chaotic process is a deterministic process. Peters (1991a) stated that a fractal shape retains its dimensionality when it is placed in an embedding dimension that is greater than its fractal dimension. This means that as the embedding dimension increase, the fractal dimension of a chaotic process will not increase as the embedding dimension.

Peters (1991b) indicated that the fractal dimension that may be calculated by the means of correlation dimension measures how an attractor fills its space. Consequently, a white-noise process is completely disorder and thus its dimension is an infinite number. On the other hand, the dimension of a chaotic system is a positive finite number (Frank *et al.*, 1988) and it need not be an integer. Liu, Granger and Heller (1993) have proved that for a true chaotic system, the correlation dimension has a stabilized value, 1, as the embedding dimension increases.

Given an embedded time series $\{x_i^m\} = [x_1^m, x_2^m, x_3^m, \ldots x_{N-m}^m]$, compute the correlation integrals whose algorithm has been illustrated in Equations (9) and (10), for different values of ε.

There are several ways of estimating the correlation dimension ν_m. For example, Denker and Keller (1986), Scheinkman and LeBaron (1989) used the ordinary linear regression; Cutler (1991) used the generalized least-square; and Ramsey and Yuan (1989) used the random coefficient regression. The most common practice in estimating ν_m is by OLS regression analysis in which the dependent variable is the natural logarithm of correlation integral $[\ln(C_{\varepsilon,m})]$ and the independent variable is the natural logarithm of tolerance $[\ln(\varepsilon)]$. The coefficient of the independent variable in this regression model will then be the estimated correlation dimension. This regression equation is as follow:

$$\ln(C_{\varepsilon,m}) = \nu_m \ln(\varepsilon) + c\,. \tag{14}$$

Repeat the estimations of ν_m with different values of m. A plot of embedding dimension (m) against their corresponding estimated correlation dimension (ν_m) is then constructed. The values on the x-axis are the values of embedding dimension and the values on the y-axis are the values of estimated correlation dimensions (ν_m).

If chaos is present in the time-series, for increasingly larger values of embedding dimension (m), the estimated correlation dimension (ν_m) will stabilize at some value. If this stabilization does not occur, it implies that ν_m increases without bounds as m increases, the system is stochastic rather than chaotic.

4. Data

The data set consists of daily Composite Indices natural logarithmic returns of SHSE and SZSE. The data covers a 10.25-year period from October 5, 1992 to December 31, 2002, consisting of 2,671 observations. The amount of data being used in this study is fairly small if compared with the time series used in the Natural Sciences. However, it contains almost 85% of the entire time series since the opening of the SHSE and SZSE. Actually, the lengths of data in most studies in Economics and Finance do not exceed 2,000 observations (Papaioannou and Karytinos, 1995).

The daily returns were calculated as the change in the logarithm of closing stock market indices of successive days:

$$x_t = \ln(P_t) - \ln(P_{t-1})\,. \tag{15}$$

Taking the first difference may not only ensure that our time series are stationary but also it is a common practice in standard econometric work to "whiten" a time series. Some scholars argue that such practice may destroy any delicate nonlinear structure present in the data (Chen, 1988). However, the economic time series have a problem that the physical sciences do not, which is as the economy grows stock prices grow. Peters (1989) pointed out that the stock prices thus have to be detrended in order to study the motion of stock prices, in other words, the economic growth have to be filtered out. Besides, Peters (1992) suggested using logarithmic data because they have the statistical properties as percentage change but they sum to their cumulative equivalent.

5. Empirical Results

5.1. *Descriptive statistics*

Table 1 gives a statistical description of the data used in the study. The positive values of skewness show that the distributions of the series are positively skewed, which are different from the results founded by Mandelbrot and Fama for the distributions of NYSE returns. On the other hand, the values of kurtosis are very large, same as those found by Mandelbrot (1964) and Fama (1965). Large kurtosis implies that the distributions are also "leptokurtic". Figure 1 shows the comparison of the distributions of the series with the normal distribution. The distributions of the returns are characterized by longer tails and higher peaks at the mean than those of the normal distribution, this difference indicates that there are strong departures from the normal distribution. Mandelbrot called this kind of distribution Stable Paretian.

The department from the normal distribution implies that the information given by both Chinese stock markets shows up in infrequent clump rather than in a smooth and continuous fashion. It contradicts the assumption of efficient market hypothesis.

Table 1. Summary Statistics: Daily SHSE and SZSE Composite Indices Returns

Statistics	SHSE	SZSE
Sample size (n)	2670	2670
Mean	0.000259	0.000136
Standard deviation	0.025986	0.024242
Skewness	1.629	0.878
Kurtosis	20.522	15.804
Minimum	−0.17905	−0.188833
Maximum	0.28860	0.272152

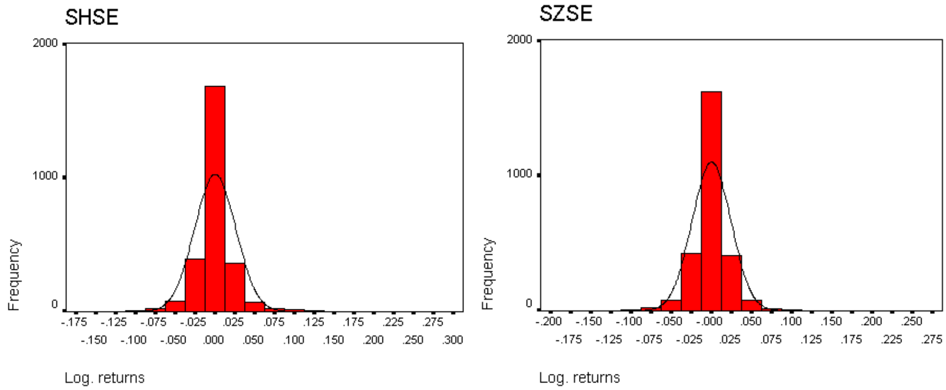

Fig. 1. The Distributions of SHSE and SZSE Composite Indices Daily Returns vs. Normal Distribution

5.2. *Rescaled range (R/S) analysis*

To test the existence of strange attractor that is characterized by fractal shapes and long-term memory in a series, it is subjected to rescaled range (R/S) analysis. Since R/S analysis is robust enough to detect long-term dependence if the distribution of the series is not normal, R/S analysis is an appropriate tool to detect long-term dependence in the data used in this study. Tables 2 and 3 present the results of the R/S analysis of the SHSE and SZSE Composite Indices daily returns.

Table 3 shows that the H exponents for daily SHSE and SZSE Composite Indices returns series are 0.617 and 0.625 respectively. The high R-square (99.6% in both) and low standard error of estimate (0.0516 and 0.0518 respectively) illustrate the goodness of fit of the regression model

Table 2. R/S Analysis Results: Daily SHSE and SZSE Composite Indices Returns

SHSE					SZSE				
n	$\ln(n)$	R/S	$\ln(R/S)$	V-statistic	n	$\ln(n)$	R/S	$\ln(R/S)$	V-statistic
10	2.3026	2.8516	1.0479	0.9018	10	2.3026	2.8506	1.0475	0.9014
20	2.9957	4.5239	1.5094	1.0116	20	2.9957	4.5348	1.5118	1.0140
25	3.2189	5.0888	1.6270	1.0178	25	3.2189	5.2138	1.6513	1.0428
40	3.6889	6.8239	1.9204	1.0790	40	3.6889	6.8821	1.9289	1.0882
50	3.9120	7.9450	2.0725	1.1236	50	3.9120	8.0276	2.0829	1.1353
100	4.6052	13.0336	2.5675	1.3034	100	4.6052	13.0982	2.5725	1.3098
125	4.8283	14.4608	2.6714	1.2934	125	4.8283	14.5323	2.6764	1.2998
200	5.2983	19.6706	2.9791	1.3909	200	5.2983	20.4227	3.0166	1.4441
250	5.5215	21.8316	3.0834	1.3808	250	5.5215	22.6745	3.1212	1.4341
500	6.2146	29.5855	3.3873	1.3231	500	6.2146	30.4383	3.4157	1.3612

Table 3. Regression Results: Daily SHSE and SZSE Composite Indices Returns

	SHSE	SZSE
Hurst exponent	0.617	0.625
P-value of the regression coefficient	0.000	0.000
Constant	−0.340	−0.358
R^2	0.996	0.996
Standard error of estimate	0.0516	0.0518

Note: Dependent variable: $\ln(R/S)_n$
Independent variable: $\ln(n)$

for estimation. The values of H are greater than 0.5 implies that persistence exists in both series. Today's data affect all future data. If the prices have been up during the current period, there is a probability of 61.7% and 62.5% respectively that they are likely to be up during the subsequent period. Such kind of persistent trend is said to be biased random process, or fractional Brownian motion. The time series is persistent implies that the investors' interpretation of events is not reflected in the price immediately. The interpretation manifests itself and becomes a bias in return, which is different from that suggested by Efficient Market Hypothesis. However, the values of H exponents are relatively low in our time series, which indicates that there may be some "white noise" and the persistent trends are not strictly consistent.

R/S analysis can be used to determine the cyclic characteristics of a time series. Figure 2 shows a plot of V-statistic versus $\ln(n)$, the cycle length can be discerned from the "break-point" in the occurring plot. Figure 2 shows that there is a "break-point" when the value of $\ln(n)$ approaches 5.3 in both series. The V-statistic flattens out after 200 days [$\ln(n) = 5.3 \rightarrow n \approx 200$]. This suggests that the series show persistence up to about 200 days and then alternate, the "memory effect" dissipates after 200 days. This implies that any trading rules or models should not use a period longer than 200 trading days, i.e. about 10 months.

5.3. *Tests of nonlinearity*: *BDS test*

To test the presence of non-linearity, the logarithmic index returns series in both stock markets are subject to BDS test. The linear dependence in the original index time series has already been removed by taking a first difference of natural logarithm. The hypotheses of the BDS test are:

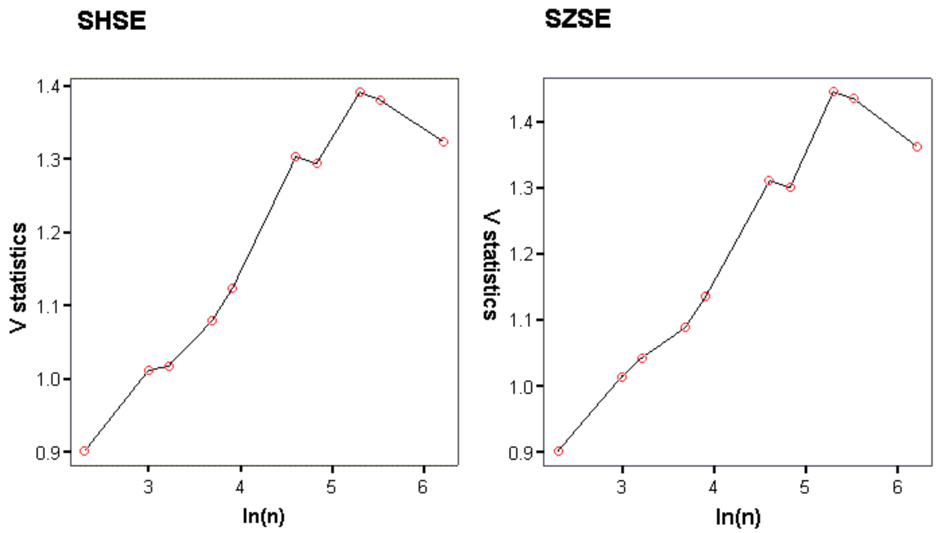

Fig. 2. The *V-Statistic* versus ln(n) Plots

H_0: The data in each time series is independent and identically distributed (I.I.D.).

H_1: Unspecified alternative (since the linear dependence has been removed, nonlinear but unspecified dependence may be found in the data).

A level of significance (α) of 5% is taken and thus the critical values for the test are ± 1.96, we should reject the null hypothesis if the BDS test statistic is greater than 1.96 or less than -1.96. We may use the standard normal distribution for the test statistic because the number of observations is more than 500, the embedding dimensions are between 2 and 5, and the ratios of ε/σ are between 0.5 and 2.

Besides the logarithmic index returns series are subjected to the BDS test, we also test the random data for I.I.D., for the purpose of comparison. In addition, we have shuffled the observations in the logarithmic index returns series that will destroy any underlying serial dependence. If the BDS test statistic for the shuffled data is not significant and indicate that the shuffled data is I.I.D., it will imply that the rejection of the null hypothesis of I.I.D. for logarithmic returns series is due to the underlying dependence in the series. Table 4 shows the test statistic in the BDS test for our returns series, the random data and the shuffled data.

Table 4 indicates that all the test statistic of logarithmic index returns series are greater than the critical value of 1.96 significantly under different

Table 4. BDS Test Results: Daily SHSE and SZSE Composite Indices Returns

		BDS test statistic				
		Log. returns series		Shuffled returns series		
ε/σ	Embedding dimension (m)	SHSE	SZSE	SHSE	SZSE	Random data
2	2	12.6805	14.6194	0.6818	−0.0664	−0.7186
2	3	16.1735	17.7598	0.5435	0.1476	−0.8881
2	4	18.5430	19.5409	0.2710	0.0305	−0.7828
2	5	19.7642	20.6998	−0.2430	−0.1819	−0.5333
1.5	2	14.6232	16.5153	0.7019	−0.0592	0.3096
1.5	3	18.2899	20.2413	0.3936	0.1477	0.2959
1.5	4	21.0134	22.9594	0.1448	−0.1894	0.2906
1.5	5	22.5724	24.9711	−0.3436	−0.4603	0.2850
1	2	15.6537	18.3951	0.3445	0.2151	−0.2560
1	3	19.8925	23.7275	0.0614	0.1242	−0.3830
1	4	23.5371	28.4080	0.2195	−0.4199	0.0279
1	5	26.5299	32.5171	−0.0482	−0.7459	0.2993
0.5	2	15.7445	19.3805	0.2158	0.4033	−0.6071
0.5	3	21.0194	27.0172	0.0360	0.1849	0.5581
0.5	4	27.3080	36.2449	0.2601	−0.2650	0.4888
0.5	5	34.0533	47.0825	−0.0457	−0.2324	−0.4122

embedding dimension and ratio of tolerance to standard deviation. Thus, we should reject the null hypothesis of I.I.D. for both series. The results strongly suggest that both series are non-linearly dependent at the 5% level of significance.

The findings of nonlinear serial correlation in both series are consistent with the other studies that the BDS test has been applied to detect nonlinear structure in the financial data. These studies include foreign exchange rate data (Hsieh, 1989), NYSE weekly stock return (Scheinkman and LeBaron, 1989), and U.S. daily return of futures contracts (Yang and Brorsen, 1993).

In addition, the findings point out that the BDS test statistic for SZSE are more than the test statistic for SHSE in most cases. It indicates that the serial correlation in SZSE is slightly more significant than that in SHSE.

On the other hand, the BDS test statistic of the simulated random data is between the critical values of −1.96 and 1.96; thus we cannot reject the null hypothesis of I.I.D. The decision of not rejecting null hypothesis indicates that there is no other dependence inside the random series.

We have shuffled the logarithmic index returns series 10 times and subjected every shuffled series to the BDS test. No results are significant at 5% level of significance. Table 4 only presents one of the results. Before shuffling the time series, the results of BDS test show that the time series are nonlinearly dependent. After shuffling the time series, the results of BDS test show that the time series are I.I.D. It implies that the underlying serial dependence has been destroyed, and the destroyed structure is nonlinearly dependent since we have filtered the linear dependence by taking a first difference of logarithms. The decision of not rejecting the null hypothesis of I.I.D. for shuffled data provides us with a strong evidence of nonlinear dependence in the original logarithmic index returns time series in both stock markets.

5.4. *Tests of chaos: correlation dimension estimation*

To test the presence of chaos, we have estimated the correlation dimensions of both series under different embedding dimensions $(m) = 1, 2, \ldots, 20$, with the delay time $(\tau) = 1$. The correlation dimensions (ν_m) are estimated by measuring the slope of $\log(C_{\varepsilon,m})$ versus $\log(\varepsilon)$. Statistically, the correlation dimension is the coefficient of the simple linear regression model, in which the dependent variable is $\log(C_{\varepsilon,m})$ and the independent variable is $\log(\varepsilon)$. Table 5 presents the estimated correlation dimensions (ν_m) of our returns

Table 5. Estimated Correlation Dimension

	Logarithmic index returns series		
Embedding dimension (m)	SHSE	SZSE	Random data
1	1.1535	2.6690	0.94
2	1.3665	3.0162	1.86
3	1.4477	3.1285	2.80
4	1.4905	3.1815	3.91
5	1.5186	3.2122	4.95
6	1.5388	3.2318	5.83
7	1.5540	3.2447	6.92
8	1.5661	3.2539	7.91
9	1.5764	3.2605	8.90
10	1.5849	3.2682	9.92
11	1.5920	3.2702	10.90
12	1.5980	3.2714	11.88
13	1.6031	3.2718	12.91
14	1.6076	3.2719	13.90
15	1.6117	3.2720	14.92
16	1.6152	3.2720	15.86
17	1.6183	3.2720	16.88
18	1.6211	3.2720	17.90
19	1.6235	3.2720	18.89
20	1.6257	3.2720	19.91

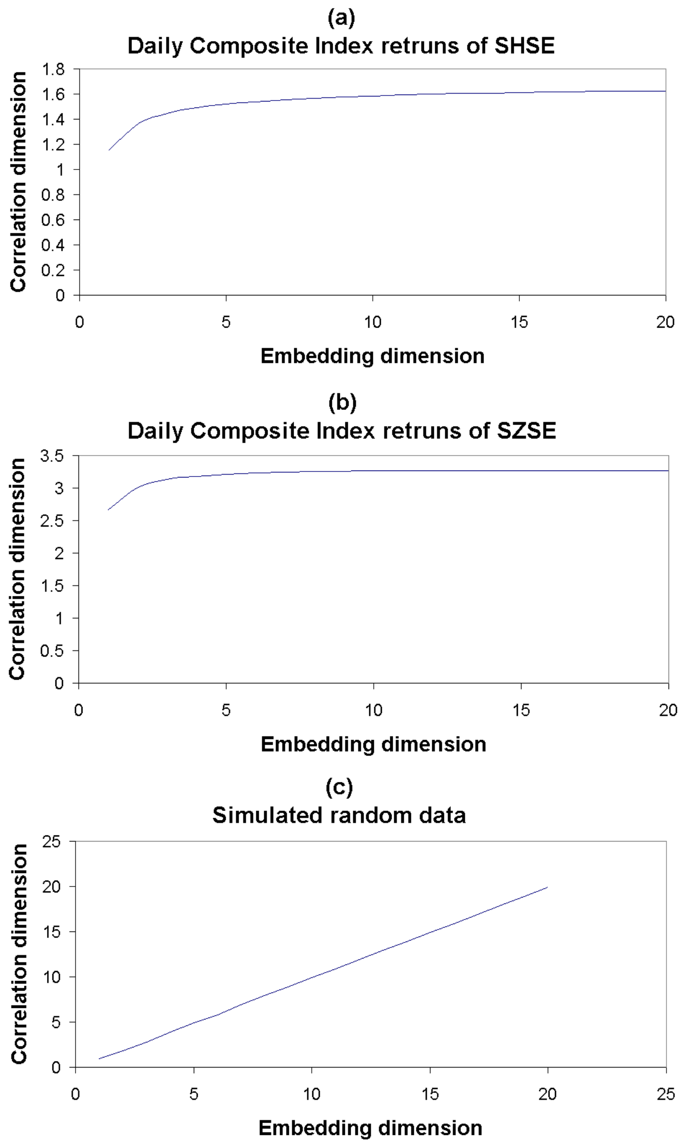

Fig. 3. The Correlation Dimension (ν) versus Embedding Dimension (m)

series at different values of embedding dimension (m). Figure 3 presents the plots of the estimated correlation dimensions (ν_m) against their respective values of embedding dimension (m) for both the series and the random data.

Figures 3(a) and (b) show that there are three distinct regimes. For low values of embedding dimension ($m < 5$), the value of correlation dimension increases rapidly as the value of embedding dimension increases. In this regime, the noise masks the indication of chaos. However, the value of embedding dimension is still low and it is very hard to conclude that the underlying system is a stochastic system.

As the value of embedding dimension increases ($5 < m < 10$), there is a weak tendency to saturate at a more or less constant value of 1.62 and 3.27 for SHSE and SZSE logarithmic index returns series respectively. This behavior indicates that our system may be chaotic.

In the third regime ($10 < m < 20$), the value of embedding dimension has been increased to 20 but the value of correlation dimension in each series does not follow the increases in the value of embedding dimension. The slope of the straight line became very small and the saturation effect has finished. For SHSE logarithmic index returns series, the saturated value is about 1.625. For SZSE, it is about 3.27.

The convergence of the estimated correlation dimensions indicate that the underlying system of both series is not a random process. Instead, it is a deterministic system. As the saturated values of correlation dimensions are all non-integers, the strange attractor is present in our series. We may conclude that our series is not only deterministic but also chaotic, which is indicated by low fractional Hausdorff dimension due to the existence of chaotic attractor.

Figure 3(c) points out that for the simulated random data, the value of correlation dimension increases as the value of embedding dimension increases. This phenomenon is similar to that reported by Barnett *et al.* (1991). The correlation dimension is infinite for a random series because it is disordered and fills the whole phase space.

6. Conclusion

In this study, we have examined the behavior of the Shanghai Stock Exchange and Shenzhen Stock Exchange Composite Indices returns series using the modified rescaled range (R/S) analysis, the BDS test and the correlation dimension estimation.

The *Hurst* exponents found in the rescaled range analysis are greater than 0.5 in both series reveal that both series are persistent, fractal and self-similar with long-term memory. However, the values of the *Hurst* exponents are relatively low (around 0.62), which indicate that both series have strong noisy components. Besides searching for memory, the V-statistic is also found for both series to detect the cyclic characteristics. The V-statistic found in this study show that the long-memory components in both series are similar, about 200 trading days.

Nonlinearity was found in our time series by the BDS test. The test statistics are very significant which indicate a strong evidence of nonlinearity. These findings are similar to those found by Hsieh (1993) who found strong evidence of nonlinear dynamics in the short-term movements of asset returns in the U.S. market. Nonlinearity is one of the indications of chaotic behavior.

Testing for chaos is a rather delicate part in this study, we have applied one quantitative technique — correlation dimension estimation to determine that the time series are chaotic or stochastic. The correlation dimensions of our time series are estimated by the regression model in which the variables are $\log(C_{m,\varepsilon})$ and $\log(\varepsilon)$ under different embedding dimensions. By plotting the values of estimated correlation dimension against the respective values of embedding dimension, the values of correlation dimension started to saturate at a value around 1.62 and 3.27 for SHSE and SZSE logarithmic index returns series respectively after the embedding dimension is larger than 6. The saturation indicates that the systems of the time series are not stochastic or random, instead, they are chaotic.

The findings of nonlinearity and chaotic behavior provide us with the following implications to economists and financial analysts:

1. The distributions of both series are leptokurtic, which agree with previous findings in the literature. These findings reveal that the quantitative economic theory copes with the possible probability distribution rather than normal distribution and should be developed for the financial analysts who want to study Chinese stock market indices. In fact, the practice of assuming normal distribution is not because of fitting the data but because of enabling us to apply the classical statistical tools.
2. The presence of memory effect in both series revealed by the rescaled range analysis indicates that both series are not generated by a pure random walk model but by a biased random walk process.
3. The existence of fractal structure implies that the Elliot Wave Principles, the Golden Triangle and the Golden Pentagon, which are mostly used in technical analysis and are developed on the basis of self-similarity can be utilized when analyzing the price movements in the Chinese stock markets.
4. The nonlinearity found in both series tells us why most of the traditional econometric tools fail to model the Chinese stock market because most of them try to "whiten" data that are originally nonlinear in order to make it suitable for these linear-based tools.

5. Nonlinear dynamics makes us know much less about how the Chinese stock markets really work than we think we do. Predicting the Chinese stock price movements may not be as secure as we believe.
6. The underlying nonlinear dynamics result in the backtesting of models and performance may have little meaning for the future because the cause-and-effect relationship is no more direct. This makes some fundamental analysis which uses the company's previous historical data less useful than expected. Such techniques include, for example, growth-stock approach, undervalued stock approach, and small capitalization approach.
7. The chaotic behavior found in our time series indicates that the predictions of the price movements in the Chinese stock market are very difficult by the traditional econometric methods, especially in the long-term. Such difficulty is due to one of the characteristics of chaotic behavior — sensitivity to initial conditions, insignificant inputs may be compounded over time and greatly influence the behavior of the system. However, prediction is possible in the short-term, before the butterfly effect dominates the system since the chaotic system is a deterministic system.

   The limitation in long-term forecasting ability also tells us why the number in the Fibonacci series greater than 377 always loses the prediction power in the reversal time and the reversal price.
8. The limitations of long-term prediction strongly suggest that the techniques of charting (technical analysis) which try to infer the short-term future behavior of the asset prices based on observed up and downturns in the past, may be applied when analyzing the price movements in the Chinese stock markets.
9. Although we cannot conclusively demonstrate that the underlying structure of the Chinese stock market is a chaotic system, both booms in May 1992 and 1996 and crashes in 1994 and 1998, periods of stability in 1995 and jarring transitions behaviors in 1998 exhibited in both SHSE and SZSE, all of these show that our time series have the characteristics of a chaotic system (Connelly, 1996).

Meanwhile, the application of this new science, chaos theory, is still in its infancy. There are few "economic chaologists" in comparison with the numbers of technical analysts and fundamental analysts. In this research, we only study the stock market index returns. Besides the stock market, some empirical findings also found that nonlinear dynamics and chaotic behaviors are present in the commodity futures markets. We may perform a similar research such as this on the series in the Chinese commodity futures markets situated in Shanghai, Dalian and Chengzhou. There is also ample room for empirical research on the capital markets in the ASEAN countries.

Besides chaos, there are two other possible explanations for the exceptional findings found in the financial and economic time series. The January 1991 issue of *Scientific American* proposed two alternatives, Wavelet Theory and Self-organized Criticality. These two alternatives are related to the chaos theory. Since they are still relatively new, there are many areas for the researchers to find some empirical evidence to support them.

In this new science, much work remains to be done.

References

Barnett, W., and M. J. Hinich, "Has chaos been discovered with economic data?" Working Paper #153, University of Texas, Austin, 1991.

Benhabib, J., and K. Nishimura, "The Hopf Bifurcation and the Existence and Stability of Closed Orbits in Multisector Models of Economic Growth". *Journal of Economic Theory* 21, 421–444 (1979).

Brock, W. A., W. Dechert, and J. Scheinkman, "A Test for Independence Based on the Correlation Dimension". Working Paper, University of Winconsin at Madison, University of Houston, and University of Chicago, 1987.

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

Chen, P., "Empirical and Theoretical Evidence of Economic Chaos". *System Dynamics Review* 4, 81–108 (1988).

Connelly, T. J., "Chaos Theory and the Financial Markets". *Journal of Financial Planning* 96(12), 26–30 (1996).

Cutler, C., "Some Results on the Behavior and Estimation of the Fractal Dimensions of Distributions on Attractors". *Journal of Statistical Physics* 62, 651–708 (1991).

Denker, G., and G. Keller, "Rigorous Statistical Procedures for Data from Dynamical Systems". *Journal of Statistical Physics* 44, 67–93 (1986).

Errunza, V., K. Hogan, O. Kini, and P. Padmanbhan, "Conditional Heteroscedasticity and Global Stock Return Distributions". *Financial Review* 29(3): 187–203 (1994).

Fama, E. F., "The Behavior of Stock Market Prices". *Journal of Business* 38 (1965).

Fama, E. F., "Efficient Capital Markets: A Review of Theory and Empirical Work". *Journal of Finance* 25, 383–417 (1970).

Frank, M. Z., R. Gencay, and T. Stengos, "International Chaos?" *European Economic Review* 32, 1569–1584 (1988).

Greene, M. T., and B. D. Fieltz, "Long Term Dependence in Common Stock Returns". *Journal of Financial Economics* 4, 249–339 (1977).

Hamill, P., and K. K. Opong, "Nonlinear Determinism in the Irish Equity Market". Unpublished Paper, School of Management, University of Ulster, Newtonabbey, 1997.

Hampton, J., "Rescaled Range Analysis: Approaches for the Financial Practitioner, Part 3". *Neuro Vest Journal* (July/August, 1996), pp. 27–30 (1996).

Howe, J. S., W. Martin, and B. Wood Jr., "Fractal Structure in the Pacific Rim". Southwestern Finance Annual Meeting, New Orleans, March, 1997.

Hsieh, D. A., "Testing for Nonlinearity in Daily Foreign Exchange Rate Changes". *Journal of Business* 62, 339–368 (1989).

Hsieh, D. A., "Chaos and Nonlinear Dynamics: Application to Financial Markets". *The Journal of Finance* 46(5), 1839–1877 (1991).

Hsieh, D. A., "Implications of Nonlinear Dynamics for Financial Risk Management". *Journal of Financial and Quantitative Analysis* 28(1), 41–64 (1993).

Huo, X. W., "Market Efficiency and Managerial Structure: Developed Markets, Emerging Markets and the Case of China". *The Progressing Chinese Stock Markets*, World Publishing Co., 1996. (The original article is written in Chinese.)

Lin, K., "The ABC's of BDS". *Journal of Computational Intelligence in Finance* (July/August, 1997), pp. 23–26 (1997).

Liu, T., C. W. J. Granger, and W. P. Heller, "Using the Correlation Exponent to Decide Whether an Economic Series is Chaotic". *Nonlinear Dynamics, Chaos and Econometrics*, John Wiley & Sons, 1993.

Mandelbrot, B., "The Variation of Certain Speculative Prices". *The Random Character of Stock Prices*, MIT Press, 1964.

Martin, V. L. and K. Sawyer, "Statistical Techniques for Modeling Nonlinearities". *Chaos and Nonlinear Models in Economics*, Edward Elgar, pp. 113–126, 1994.

Opong, K. K., G. Mulholland, A. F. Fox, and K. Farahmand, "The Behavior of Some UK Equity Indices: An Application of Hurst and BDS Tests". *Journal of Empirical Finance* 6(3): 267–282 (1999).

Pandey, V., T. Kohers, and G. Kohers, "Nonlinear Determinism in the Equity Markets of Major Pacific Rim Countries and the United States". Southern Finance Association Conference, New Orleans, LA, 1997.

Papaioannou, G. and A. Karytinos, "Nonlinear Time Series Analysis of the Stock Exchange: The Case of an Emerging Market". *International Journal of Bifurcation and Chaos* 5(6), 1557–1584 (1995).

Peters, E. E., "Fractal Structure in the Capital Markets". *Financial Analyst Journal* (July/August, 1980), pp. 32–37 (1989).

Peters, E. E., *Chaos and Order in the Capital Markets*. John Wiley & Sons, 1991.

Peters, E. E., "A Chaotic Attractor for the S&P 500". *Financial Analyst Journal* (March/April, 1991), pp. 55–62 (1991).

Peters, E. E., "R/S Analysis Using Logarithmic Returns". *Financial Analyst Journal* (November/December, 1992), pp. 81–82 (1992).

Peters, E. E., *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*. John Wiley & Sons, 1994.

Ramsey, J., and H. Yuan, "Bias and Error Bars in Dimension Calculations and Their Evaluation in Some Simple Models". *Physics Letters* A-134, 287–297 (1989).

Sakai, H., and H. Tokumaru, "Autocorrelations of a Certain Chaos". *IEEE Transactions on Acoustics, Speech, and Signal Processing* 28(5), 588–590 (1980).

Scheinkman, J., and B. LeBaron, "Nonlinear Dynamics and Stock Returns". *Journal of Business* 62, 311–337 (1989).

Takens, F., *Lecture Notes in Mathematics No. 366*. Springer, 1981.

Willey, T., "Testing for Nonlinear Dependence in Daily Stock Indices". *Journal of Economics and Business* 44(1), 63–76 (1992).

Yang, S. R., and B. W. Brorsen, "Nonlinear Dynamics of Daily Futures Prices: Conditional Heteroskedasticity or Chaos?" *The Journal of Futures Markets* 13(2), 175–191 (1993).