Chaotic behavior in national stock market indices
New evidence from the close returns test

Michael D. McKenzie*

School of Economics and Finance, RMIT University, GPO Box 2476V, Melbourne, Victoria 3001, Australia
Received 30 June 1999; received in revised form 27 March 2000; accepted 5 May 2000

Abstract

Attempts have been made to detect chaotic behaviour in financial markets data using techniques which require large, clean data sets. Although such data are common in the physical sciences where these tests were developed, financial returns data typically do not conform. The close returns test is a recent innovation in the literature and is better suited to testing for chaos in financial markets. This paper tests for the presence of chaos in a wide range of major national stock market indices using the close returns test. The results indicate that the data are not chaotic, although considerable nonlinearities are present. The commonly used BDS test is also applied to the data and, in comparison, the close returns test provides substantially more evidence of nonlinearity compared to the BDS test. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Nonlinearity; Chaos; BDS; Close returns test

1. Introduction

The discovery, that deterministic nonlinear equations could generate data which appear random, provided a major breakthrough in the way scientists viewed a wide range of physical processes and natural phenomena. While by no means a complete list, chaos has been identified in hydrodynamic turbulence, lasers, electrical circuits, chemical reactions, disease

* Level 16, 239 Bourke Street, Melbourne, Victoria 3000, Australia. Tel.: +61-3-9925-5891; fax: +61-3-9925-5986.
E-mail address: michael.mckenzie@rmit.edu.au (M.D. McKenzie).

1044-0283/01/$ – see front matter © 2001 Elsevier Science Inc. All rights reserved.
PII: S1044-0283(01)00024-2
epidemics, biological reactions, and climatic change.\textsuperscript{1} Buoyed by the research of their physical science counterparts, financial market researchers have attempted to establish whether the apparently random nature of asset prices and economic time series could also be explained by the presence of chaotic behaviour.

One of the most commonly applied tests for nonlinearity is the BDS test of Brock, Dechert, and Scheinkman (1987), details of which may be found in Dechert (1996). Subsequent to its introduction, the BDS test has been generalised by Savit and Green (1991) and Wu, Savit, and Brock (1993) and more recently, DeLima (1998) introduced an iterative version of the BDS test. The BDS test is a statistical test of the null hypothesis of IID and is based on the Grassberger and Procacia (1983) correlation integral. As such, the BDS procedure may be considered as a test for linear and nonlinear departures from IID rather than a specific test for chaos. It is in this latter context however, that the test has most commonly been applied usually in conjunction with the estimation of entropy, Lyapunov exponents, or correlation dimensions.

The BDS test has been used to test for nonlinear behaviour in a wide range of financial data including national stock market indices (see Abhyankar, Copeland, & Wong, 1995, 1997; Ahmed, Rosser, & Uppal, 1996; Barkoulas & Travlos, 1998; Hsieh, 1991; Mayfield & Mizrach, 1992; Olmeda & Perez, 1995; Philippatos, Pilarinu, & Malliaris, 1993; Scheinkman & LeBaron, 1989; Sewell, Stansell, Lee, & Below, 1996; Willey, 1992), exchange rates (see Cecen & Erkal, 1996; Chiarella, Peat, & Stevenson, 1994; Hsieh, 1989; Serletis & Gogas, 1997; Vassilicos, Demos, & Tata, 1992), futures data (see Chwee, 1998; Eldridge & Coleman, 1993; Kodres & Papell, 1991; Vaidyanathan & Krehbiel, 1992), and commodity prices (see Frank & Stengos, 1988; Kohzadi & Boyd, 1995). In general, the BDS test results furnished by this literature provide substantial empirical evidence of nonlinear structure in a wide range of financial asset prices.

The BDS test belongs to the metric invariant class of tests for chaotic behaviour, which were developed for application in the physical sciences where long, clean data sets are the norm. In finance, however, small noisy data sets are more common and the application of the BDS test to such data presents a number of problems (a discussion of which is presented in Section 2). A recent development in the literature has been the introduction of the close returns test, which is a topological invariant testing procedure.\textsuperscript{2} Compared to the existing metric class of testing procedures including the BDS test, the close returns test is better suited to testing for chaos in financial and economic time series. Despite these advantages, in comparison to the BDS test, the close returns test has largely been overlooked in the finance literature. Gilmore (1993a, 1993b) applied the close returns test to weekly CRSP data sampled over the period 1962–1989 and the results suggest the presence of nonlinear structure in the data, which was not chaotic. US macroeconomic, treasury bill, and exchange rate data were also considered with similar results. Gilmore (1996) extended the earlier

\textsuperscript{1} See Gleick (1987) for discussion on the range of processes, which have been found to exhibit chaotic behaviour.

\textsuperscript{2} See Gilmore (1998) for a survey.
analysis of the CRSP data set and found that linear filters and GARCH modelling did not fully capture all of the structure in the data. Gilmore (2000) used the close returns test to test for the presence of chaos in daily foreign exchange returns and while no evidence of chaos was found, the data did exhibit nonlinearity.

The purpose of this article is to apply the close returns test for low-dimensional chaos to daily data sampled from 12 national stock market indices. To allow comparisons to be made, the BDS test will also be applied to these data. Thus, this study will augment the evidence on nonlinearity in stock markets provided by Gilmore (1993a, 1993b, 1996) for CRSP data to encompass a wider range of markets. The application of the BDS test to the same data will allow comparisons to be drawn between the results derived using a topological and a metric method. Further, as some of the markets included in this study have been tested previously in the literature using the BDS methodology, comparisons may also be drawn to these prior results. For example, Willey (1992) examines daily S&P 100 and NASDAC index returns sampled over the period 1982–1988 using the BDS statistics and concludes “... no underlying dependence, chaotic or otherwise, is present in the data” (p. 72). In contrast, Sewell et al. (1996) reject the null of IID for stock market indices for Japan, Hong Kong, Singapore, and the S&P 500 using weekly data sampled over the period 1980–1994. Abhyanark et al. (1997) strongly reject IID behaviour in real-time returns data for the S&P 500, NIKKEI, DAX, and FTSE sampled over the period September–November 1991. The rest of this paper proceeds as follows. Sections 2 and 3 introduce and discuss the limitations and prospects of the BDS and close returns test, respectively. Section 4 formally introduces the close returns test procedure for low-dimensional chaos. Section 5 discusses the data to be tested and presents the estimation results. The BDS test is initially applied to the data followed by the close returns test and comparisons drawn. Finally, Section 6 presents some concluding comments and discusses areas for further research.

2. The BDS test

In general, the literature of BDS testing of financial asset prices has provided substantial empirical evidence of the presence of nonlinear structure. One must be careful when interpreting this evidence, however, as there are a number of problems associated with the estimation of the BDS test. At a general level, such criticisms include the following. First, the BDS test is not trivial to calculate. Second, the threshold term (ε) and the embedding dimension (M) in the BDS test are subjective and it is not clear how to interpret the results when their significance depends on the values assigned to ε and M. Third, no formal distribution theory exists for the correlation dimension, which makes precise statistical testing

---

3 Hsieh (1991) also studied CRSP returns and similarly found that the standard GARCH model did not fully capture all of the nonlinearity present in the data.

4 LeBaron (1997) has made advances in this area.

5 Dechert (1995) proposes a theoretical procedure for selecting ε in the BDS test.
difficult. Fourth, Grassberger (1990) argues that dimension estimates, which are not corrected for dynamic correlation, may possibly produce spurious evidence of low-dimensional chaos. Only where geometric correlation is tested for may the evidence be considered robust, and failure to consider this point has rendered many previously reported dimension calculations ‘obsolete.’ A fifth problem is that the time-ordering of the data is not preserved, which means that important information about the nature of the underlying process by which the data are generated is potentially discarded in the testing process.

In addition to these five general problems associated with the application of the BDS test, there also exist a number of additional concerns in the current context of testing financial data. The first problem is that noise may render any dimension calculation useless (Brock, 1986). While this problem applies to all data, it is especially relevant when testing financial data, which typically contain a relatively large amount of noise, and the noise-to-signal ratio increases as the sampling interval shortens. A second problem is that the BDS test is not robust to the presence of linear relationships, which are common to financial data, and so necessitates prewhitening using an autoregressive filter. Brock and Sayers (1988) have shown that the power of the BDS test declines as the linear filter used to prewhiten the series becomes more complex, which, Willey (1992, p. 67) argues, necessitates setting the lag length in the AR model as low as possible. Hence, a trade-off exists as it is necessary to eliminate as much of the linear dependence as possible to lessen the chance of generating spurious results, yet the order of the AR model needs to be set as low as possible to increase the power of the BDS test.

The third problem relates to the observation that the correlation dimension converges to the dimension of the chaotic attractor as the number of observations approaches infinity. As such, the required data sample size quickly increases with the attractor dimension and the use of relatively small data sets may lead to spurious results resulting from substantial bias (see Ramsay & Yuan, 1989, 1990; Ruelle, 1991). Although the search in finance is for low-dimensional chaos, sufficiently large data sets are nonetheless unlikely especially for economic data, which typically provide only quarterly or monthly observations at best. Thus, while the BDS test has proven to be a popular test for nonlinearity, it does possess a number of problems, which limit the extent to which we may reliably interpret the results.

3. Topological tests of chaos

A recent development in the literature has been the development of topological invariant testing procedures (see Gilmore, 1998 for a survey). The topological approach to testing for

---

6 Brock and Baek (1991) contain a discussion of how standard errors can be estimated for the correlation dimension.

7 Wolf, Brandstater, and Swift (1985) specify the minimum number of observations as \( N = a^D \) where \( N \) is the number of observations, \( D \) is the attractor dimension, and \( a \) is on the order of 10.

8 Science is unlikely to uncover high-dimensional chaotic processes as computers have only finite resolution and data sets have only finite length in the real world.
chaos has origins as far back as Poincaré (1892) and attempts to determine how the unstable periodic orbits of the strange attractor are intertwined. The processes of stretching and compression are responsible for organising the strange attractor in a unique way and if one can determine how the unstable periodic orbits are organised, we can identify the stretching and compressing mechanisms responsible for the creation of the strange attractor. This information is robust against noise and is independently verifiable. Once these mechanisms have been identified, a geometric model can be constructed, which describes how to model the stretching and squeezing mechanisms responsible for generating the original time series. That is to say, topological tests may not only detect the presence of chaos (the only information provided by the metric class of tests), but can also provide information about the underlying system responsible for the chaotic behaviour. As the topological method preserves the time ordering of the data, where evidence of chaos is found, the researcher may proceed to characterise the underlying process in a quantitative way. Thus, one is able to reconstruct the stretching and compressing mechanisms responsible for generating the strange attractor. While not enabling the researcher to identify the underlying equation system, it does allow the rejection of models, which produce behaviour that is incompatible with the characteristics of the strange attractor identified by this technique.

The close returns test is a topological-based testing procedure, which was specifically designed to detect low-dimensional chaotic behaviour. It is a two-part test consisting of a qualitative component, which is a graphical test for the presence of chaotic behaviour. The second quantitative element is a test of the null hypothesis that the data are IID against both linear and nonlinear alternatives. Full details of this testing procedure may be found in Section 4. Compared to the existing metric class of testing procedures, the close returns test is better suited to testing for chaos in financial and economic time series. First, unlike the BDS test, which only tests for nonlinearity (and relies on other metric tests for evidence of chaos), the close returns test provides evidence as to departures from IID as well as the presence of chaotic behaviour. Second, smaller data sets will produce evidence of chaotic behaviour, although the pattern of cycling around the chaotic attractor’s unstable periodic orbit will be more visible for larger data sets. Third, test results for noisy data sets will still produce results, which indicate chaos where present. Fourth, linear dependence in the data is not a problem for the close returns test, as a series exhibiting autoregressive elements will generate results different from that of a chaotic series. Fifth, as the close returns plots are only the first of a six-step test process, in the event that chaotic behaviour is erroneously found in the data, the chaotic explanation will be correctly rejected in the subsequent steps.

---

9 Poincaré (1892) first proposed studying differential equations by examining how an entire neighbourhood in phase space evolved.

10 Mindlin, Hou, Solari, Gilmore, and Tuñillaro (1990) have theoretically shown how one may reconstruct the strange attractor based on the topological test and Mindlin, Solari, Natiello, Gilmore, and Hou (1991) reconstruct the underlying dynamic mechanism for experimental data.

11 Mindlin and Gilmore (1992, p. 232) demonstrate that in data with an equal signal-to-noise ratio, the signal in the data can be recovered using a low-pass filter such as a moving average.
While the close returns test does have a number of distinct advantages over the metric invariant class of tests including the BDS test, it is not without some problems of its own. First, the threshold term in the close returns test is subjective. Unlike the BDS test, however, this is not a significant problem as the threshold term can be varied without altering the qualitative nature of the pattern of close returns observed. Thus, the estimation results should not be critically dependent on the value chosen for $\varepsilon$ (a discussion on this point may be found in Section 4). Second, both the metric and topological tests suffer a problem where the data are nonstationary. For example, regime shifts may either cause a temporary disturbance to the path of the data or permanently alter the strange attractor such that it takes on a new form. This is especially a concern when testing data for emerging countries, which typically undergo significant regulatory changes in the development period (e.g., Sewell et al., 1996 included Korean data; Ahmed et al., 1996 tested Pakistan stock market data; Serletis & Gogas, 1997 tested black market foreign exchange rates for a number of East European Countries; and Barkoulas & Travlos, 1998 tested Greek stock market data). While nonstationarity is a problem for both the topological and metric classes of tests, it is more severe for the latter, as it does not maintain the temporal ordering of the data. As the close returns test searches over relatively nearby points in the time series, the presence of a trend (which obscures widely separated segments) does not necessarily obscure the relationship between nearby segments.

4. The close returns test

Following Gilmore (2000), the close returns test searches for unstable periodic orbits embedded in the strange attractor. This means that as the chaotic series $\{x_t\}$ evolves if an observation, $x_t$, occurs near a periodic orbit, then subsequent observations will return to the neighbourhood of $x_t$ after some interval ($I$). In this case, $I$ indicates the length of the orbit measured in units of the sampling rate. Thus, the difference between $x_t$ and $x_{t+I}$ will be arbitrarily small although nonzero. Further, if the observations evolve near the periodic orbit for a sufficiently long time, we will also observe $x_{t+1}$ near $x_{t+1+I}$, $x_{t+2}$ near $x_{t+2+I}$, and so on, over a number of subsequent observations. The close returns test attempts to identify such segments of the data in which the differences $|x_t - x_{t+I}|$ are small.

To detect these regions of close returns in the data, all differences $|x_t - x_{t+I}|$ are computed, where $t=(1 \ldots n)$, $i=(1 \ldots n-1)$, and $n$ is the sample length. To determine whether these differences are significantly small, some threshold value, $\varepsilon$, is arbitrarily chosen. Typically, $\varepsilon$ is set at 2–5% of the maximum difference between any two values in the data set and in this study, 5% is used. The results of this procedure are not critically dependent on the value chosen for $\varepsilon$ as this threshold term can be varied without altering the qualitative nature of the pattern of close returns observed in the data. Where a value of $\varepsilon$ is chosen, which is too small, the signal will be less obvious due to a lack of observations and too large a value of $\varepsilon$ will serve to obscure the signal. Ideally, $\varepsilon$ should be set so as to generate a sufficient array of points to allow the qualitative nature of any period of close returns to be identified.

The close returns test consists of two parts. The first part of the test is qualitative and consists of creating a graph, which summarises this close returns information. The $x$-axis
of the graph refers to the values of \( t \), and the y-axis of the figure refers to values of \( i \). If \(|x_t - x_{t+1}| < \varepsilon\), the result is coded black; if \(|x_t - x_{t+1}| > \varepsilon\), the result is coded white. As shown in Gilmore (1993a), if the time series is IID, the distribution of the black dots shall be random with no pattern evident. If the time series exhibits chaotic behaviour, however, the close returns caused by the almost periodic behaviour of the strange attractor shall manifest itself as horizontal line segments in the plot. These horizontal lines indicate that the differences between observations are very small over segments of the time series. For example, Fig. 1 depicts a region of a close returns plot of a 2000 observation data series generated by the Logistic map, which is a univariate chaotic system of the form:

\[
x_t = Ax_{t-1}(1 - x_{t-1}),
\]

where \( A = 3.6 \) and \( x \) is a random variable \( 0 < x < 1 \). The close returns plot generated by this series clearly contains horizontal line segments indicating the true chaotic nature of this set of observations.

The second part of the close returns test for chaos consists of constructing a histogram, which summarises the information of the close returns plot. This histogram records the

---

12 The close returns plot for chaotic data generated by a Hénon map, as well as a random series, may be found in Gilmore (2001).
number of close returns for each \( i \), where \( H_i = \sum \Theta(\varepsilon - |x_i - x_{i+1}|) \) and \( \Theta \) is the Heaviside theta function, i.e., \( \Theta(x) = +1 \) if \( x > 0 \) and \( \Theta(x) = -1 \) if \( x < 0 \). Thus, the histogram presents information about the number of close returns \((H_i)\) for each \( i \). As shown in Gilmore (1993a), if the series is chaotic, the histogram will contain a series of peaks. If the series is IID, however, the histogram will be uniformly distributed around some value \( \bar{H} \). To determine whether the null of IID can be accepted for the data, we need to establish whether \( H_i = \bar{H} \) on average. This may be tested by calculating the sample chi-squared test statistic \((\chi^2_i)\), which for this binomial distribution may be expressed as:

\[
\chi^2_i = \frac{\sum_{i=1}^{k} [H_i - \bar{H}]^2}{np},
\]

where \( \bar{H} \) is equal to the number of observations over which the number of close returns is counted \((n)\) times the probability of a hit \((p)\), i.e., \( \bar{H} = np \), where \( p \) can be estimated as [Eq. (2)]:

\[
p = \frac{\text{total number of close returns}}{\text{total area of plot}}.
\]

The estimated \( \chi^2_i \) is then compared to the critical chi-squared test value \((\chi^2_c)\), which is estimated with \((k-1)\) degrees of freedom. The decision rule to be applied is if the \( P \) value generated by this test is less than .05, the null hypothesis that the data are IID is rejected.

This \( \chi^2 \) test applied to the histogram generated by a close returns plot is specifically designed to emphasise the recurrence property common to chaotic data. If nonlinearity is present, which is not chaotic however, this histogram test is not well equipped to detect its presence. Thus, an alternative quantitative box-plot test may be specified, which attempts to identify general nonlinearity in the data. This test applies the \( \chi^2 \) test directly to the close returns plot itself. The box-plot test consists of dividing the close returns plot into boxes of \( n \times n \) observations and \( H_i \) is estimated for each \( n \times n \) box. \( \bar{H} \) is estimated as \( \bar{H} = (n \times n \times p) \) and the \( \chi^2 \) test applied to the squared differences as per Eq. (1). The \( \chi^2_i \) may be compared to the estimated \( \chi^2_c \) and where the \( P \) value is less than .05, the null of IID is rejected in the data.

5. Results

5.1. Data

Daily stock market index data for 10 countries were sourced from the Datastream database for the period January 1990–December 1998, giving a total of 2349 observations. The 10 countries and their respective price indices are: Australia (All Ordinaries Accumulation Index [ALL ORDS]), Canada (Toronto Stock Exchange 300 [TSE 300]), France (CAC 40), Germany (German Aktien Index (DAX)), Hong Kong (Hang Seng),
Japan (Nikkei Dow), Singapore (Straits Times), Switzerland (SWISS MI), United Kingdom (FTSE 100), and the United States (S&P 500, S&P 100, and the NASDACL index). Three indices were chosen for the US as each has been studied in the literature and their inclusion allows comparisons to be drawn between the close returns results and that previously reported.

Each of these national stock price indices was transformed via first differencing of the log price data to create a series, which approximates the continuously compounded percentage return. Extensive stationarity testing using the Augmented Dickey Fuller and Phillips–Peron procedures indicated that each of these return series was mean-reverting. To provide a basis for comparison to the results of the close returns test, the popular BDS test for nonlinearity may be applied to this data. Prior to estimating the BDS test, it is necessary to prewhiten these returns data as the BDS test is sensitive to the presence of linear structure. Most commonly, autoregressive (AR) equations are fitted to the data, which in some cases have been augmented with dummy variables to take account of daily returns seasonality, the crash, and so on. Mindful of the evidence provided by Brock and Sayers (1988), in this paper, the mean equation was limited to a standard AR model and only significant autocorrelation for the first 10 lag intervals was included in the model. Details of the lag structure of each AR model and the estimated coefficients as well as their associated $t$ statistics are presented in Table 1. Each of the national stock market returns series exhibited significant autocorrelation, which were modelled as an AR process. For the S&P 500 index and the Hong Kong market index, the estimated AR model captured all of the linear structures present. For the remaining stock market indices, some additional residual autocorrelation was evident at lag lengths

<table>
<thead>
<tr>
<th>Index</th>
<th>Model summary</th>
<th>$R^2$</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>$R_t = 0.0002 + 0.08R_{t-1} \ (4.33) - 0.04R_{t-2} \ (2.17)$</td>
<td>0.009</td>
<td>0.008</td>
</tr>
<tr>
<td>Canada</td>
<td>$R_t = 0.0001 + 1.18R_{t-1} \ (9.01) + 0.05R_{t-3}$</td>
<td>0.040</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(2.78) - 0.04R_{t-8} \ (2.29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>$R_t = 0.0002 + 0.04R_{t-1} \ (2.05) - 0.05R_{t-7} \ (2.66)$</td>
<td>0.004</td>
<td>0.011</td>
</tr>
<tr>
<td>Germany</td>
<td>$R_t = 0.0004 - 0.05R_{t-2} \ (2.50) - 0.04R_{t-6} \ (2.38)$</td>
<td>0.005</td>
<td>0.012</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>$R_t = 0.0005 + 0.09R_{t-3} \ (4.83) - 0.04R_{t-4}</td>
<td>0.016</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(2.17) - 0.05R_{t-7} \ (2.47) + 0.04R_{t-10} \ (2.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>$R_t = -0.0004 - 0.07R_{t-2} \ (3.48) + 0.07R_{t-9} \ (3.43)$</td>
<td>0.010</td>
<td>0.014</td>
</tr>
<tr>
<td>Singapore</td>
<td>$R_t = 0.0005 + 0.18R_{t-1} \ (9.29) - 0.05R_{t-7}</td>
<td>0.040</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(2.52) + 0.04R_{t-10} \ (2.27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>$R_t = 0.0005 + 0.04R_{t-1} \ (2.23) + 0.05R_{t-4} \ (2.47)$</td>
<td>0.004</td>
<td>0.010</td>
</tr>
<tr>
<td>UK</td>
<td>$R_t = 0.0003 + 0.07R_{t-1} \ (3.56) - 0.07R_{t-7} \ (3.64)$</td>
<td>0.011</td>
<td>0.008</td>
</tr>
<tr>
<td>US S&amp;P 100</td>
<td>$R_t = 0.005 - 0.07R_{t-7} \ (3.45) + 0.05R_{t-10} \ (2.44)$</td>
<td>0.007</td>
<td>0.008</td>
</tr>
<tr>
<td>US S&amp;P 500</td>
<td>$R_t = 0.0006 - 0.04R_{t-3} \ (2.06) - 0.04R_{t-6} \ (2.33)$</td>
<td>0.011</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>-0.06R_{t-7} \ (3.33) + 0.04R_{t-10} \ (2.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US NASDAC</td>
<td>$R_t = 0.0009 - 0.07R_{t-7} \ (3.52)$</td>
<td>0.005</td>
<td>0.014</td>
</tr>
</tbody>
</table>

Note: Absolute values of $t$ statistics for lag coefficients in parentheses.
greater than 10, although this omitted structure consisted of three significant lag coefficients at most.\textsuperscript{13}

5.2. BDS test results

The BDS test may be applied to the residuals of the AR models summarised in Table 1. In estimating the BDS test, it is necessary to specify the embedding dimension and the threshold term. As a general guide, Brock, Hsieh, and LeBaron (1991) suggest an embedding dimension of two to five for data of a length such as that tested in this study. Further, the threshold term should include values ranging from 0.5 to 2. Table 2 presents the results of the BDS test estimated for each combination of embedding and threshold value. For the data in this study, which exceed 500 observations, Brock et al. (1991) and Hsieh (1991) show the sample distribution of the BDS statistic to be well approximated by an asymptotic normal distribution. Thus, a value of greater than 1.96 signifies a rejection of the null that the data are IID at the 5\% level of significance.

From Table 2, the BDS test results provided some evidence of nonlinear structure in each index except for Australia, in which none of the estimated BDS statistics is significant. For the remaining 10 indices tested, the estimated BDS test rejects the null of IID although the strength of this evidence is somewhat mixed as the chosen threshold and embedding dimension values greatly influence the estimated results. For example, the UK and French market index only generated one instance in which the estimated BDS test was significant. Similarly, all three US market indices tested only provided two test results, which rejected the null hypothesis of IID. For other market indices, however, more substantial evidence of nonlinearity is evident. The German and Swiss indices generated seven significant BDS test scores from the 16 threshold/embedding dimension combinations tested. Further, the Hong Kong market index generated eight instances where the BDS test rejected the null of IID and the Singapore market provided nine such cases. One problem with the BDS test, is that it is not immediately obvious how best to interpret such results where the rejection of the null is contingent on the values assigned to the threshold and embedding dimension.

As the data are sampled over a relatively long time period, it is possible that regime changes may have had an impact on the market. Such structural change has the potential to cause spurious BDS test results.\textsuperscript{14} Hsieh (1991, p. 428) suggests subperiod analysis as a means of determining whether regime shifts are responsible for significant BDS test results such as those found in Table 2. If the estimated results consistently provide evidence of nonlinearity across the various periods tested, Hsieh argues one can discount structural change as the cause of the nonlinearity. In light of this, each series was split in half and the BDS test applied to each subperiod. The results of this procedure are not presented to

\textsuperscript{13} The BDS test was also applied to the residuals of an AR equation, which removed all linear dependence from the data, and the results were qualitatively unchanged from those reported here.

\textsuperscript{14} DeLima (1998) introduced an iterative version of the BDS test, which allows the researcher to distinguish between rejection of the null due to nonlinearity or regime changes.
Table 2
BDS statistics for national stock market returns

<table>
<thead>
<tr>
<th>εσ</th>
<th>M</th>
<th>Australia</th>
<th>Canada</th>
<th>France</th>
<th>Germany</th>
<th>Hong Kong</th>
<th>Japan</th>
<th>Singapore</th>
<th>Switzerland</th>
<th>UK</th>
<th>US S&amp;P 100</th>
<th>US S&amp;P 500</th>
<th>US NASDAQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>2</td>
<td>.1604</td>
<td>.4115</td>
<td>.0890</td>
<td>.4102</td>
<td>.7050</td>
<td>.3920</td>
<td>.9080</td>
<td>.4410</td>
<td>.1442</td>
<td>.2637</td>
<td>.2517</td>
<td>.1794</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>.1274</td>
<td>.3689</td>
<td>.0811</td>
<td>.4176</td>
<td>.7227</td>
<td>.3774</td>
<td>.9377</td>
<td>.4345</td>
<td>.1460</td>
<td>.2435</td>
<td>.2304</td>
<td>.1774</td>
</tr>
<tr>
<td>4</td>
<td>.0732</td>
<td>.2312</td>
<td>.0516</td>
<td>.2915</td>
<td>.5335</td>
<td>.2379</td>
<td>.6776</td>
<td>.2882</td>
<td>.0919</td>
<td>.1568</td>
<td>.1505</td>
<td>.1109</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.0372</td>
<td>.1320</td>
<td>.0303</td>
<td>.1736</td>
<td>.3401</td>
<td>.1445</td>
<td>.4356</td>
<td>.1662</td>
<td>.0452</td>
<td>.0999</td>
<td>.0940</td>
<td>.0605</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>.3952</td>
<td>.3988</td>
<td>.2983</td>
<td>.9224</td>
<td>1.3144</td>
<td>.8197</td>
<td>1.5916</td>
<td>1.0130</td>
<td>.8207</td>
<td>.3703</td>
<td>.5711</td>
<td>.5099</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>.5743</td>
<td>1.4775</td>
<td>.4597</td>
<td>1.6338</td>
<td>2.2978*</td>
<td>1.3452</td>
<td>2.8380*</td>
<td>1.7243</td>
<td>.6784</td>
<td>.9273</td>
<td>.8991</td>
<td>.7786</td>
</tr>
<tr>
<td>4</td>
<td>.6010</td>
<td>1.6364</td>
<td>.5216</td>
<td>1.9791*</td>
<td>2.8160*</td>
<td>1.5007*</td>
<td>3.4303*</td>
<td>1.9743*</td>
<td>.7696</td>
<td>1.0190</td>
<td>1.0037</td>
<td>.8520</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.5480</td>
<td>1.6059</td>
<td>.5070</td>
<td>1.9845*</td>
<td>2.9372*</td>
<td>1.5248*</td>
<td>3.6021*</td>
<td>1.9747*</td>
<td>.7044</td>
<td>1.0758</td>
<td>1.0756</td>
<td>.8072</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>2</td>
<td>.4694</td>
<td>.9918</td>
<td>.4424</td>
<td>.8550</td>
<td>1.1948</td>
<td>.7767</td>
<td>1.3867</td>
<td>1.0440</td>
<td>.4700</td>
<td>.5627</td>
<td>.4877</td>
<td>.4533</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>.8861</td>
<td>1.9935</td>
<td>.8829</td>
<td>1.9777*</td>
<td>2.5406*</td>
<td>1.6023</td>
<td>3.0545*</td>
<td>2.2298*</td>
<td>1.0928</td>
<td>1.2147</td>
<td>1.1794</td>
<td>1.1657</td>
</tr>
<tr>
<td>5</td>
<td>.13957</td>
<td>3.4152*</td>
<td>1.5282</td>
<td>3.7463*</td>
<td>4.7344*</td>
<td>2.8279*</td>
<td>5.6763*</td>
<td>3.9952*</td>
<td>1.9183</td>
<td>2.1542</td>
<td>2.2050*</td>
<td>1.9888*</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>.3681</td>
<td>.7392</td>
<td>.4364</td>
<td>.5158</td>
<td>.8400</td>
<td>.5242</td>
<td>.9087</td>
<td>.7942</td>
<td>.3932</td>
<td>.4232</td>
<td>.3693</td>
<td>.3486</td>
</tr>
<tr>
<td>3</td>
<td>.7685</td>
<td>1.7035</td>
<td>1.0116</td>
<td>1.4143</td>
<td>1.9256</td>
<td>1.2505</td>
<td>2.2047*</td>
<td>1.8879</td>
<td>1.0230</td>
<td>1.0265</td>
<td>.9986</td>
<td>1.0113</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.1539</td>
<td>2.6487*</td>
<td>1.6194</td>
<td>2.4230*</td>
<td>3.0647*</td>
<td>1.9837*</td>
<td>3.5914*</td>
<td>2.9863*</td>
<td>1.7131</td>
<td>1.5793</td>
<td>1.6064</td>
<td>1.5988</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.4986</td>
<td>3.5898*</td>
<td>2.1617*</td>
<td>3.3533*</td>
<td>4.1669*</td>
<td>2.7275*</td>
<td>4.9181*</td>
<td>4.0855*</td>
<td>2.2998</td>
<td>2.1656*</td>
<td>2.2350*</td>
<td>2.1633*</td>
<td></td>
</tr>
</tbody>
</table>

The following table presents the estimated BDS statistics for a given embedding dimension (M) and threshold (ε) for various national stock market index daily returns sampled over the period January 1990–December 1998.

* Denotes significance at the 5% level.
conserve space and are available on request. In summary, the BDS test applied to data sampled in the first subperiod generated far fewer instances of a significant test result. For Australia, Canada, France, UK, and all three US indices, none of the BDS tests provided evidence of nonlinearity. Further, for the other five country indices in which evidence of nonlinearity is found, it is far more limited than was the case in the full sample estimation. The BDS test results for the second subperiod again provide less evidence of nonlinearity compared to the full sample period results. Only the test results for Singapore were qualitatively unchanged; as for the other countries, the BDS test results provided fewer instances where the null was rejected. It is difficult to objectively assess the extent to which this subperiod analysis provides evidence of significant BDS test results being driven by regime changes rather than the presence of nonlinearity. Certainly, in the case of Canada, for example, the fact that no nonlinearity was detected in the first subperiod, yet six BDS test scores were significant in the second subperiod, would tend to suggest that some form of regime change may have taken place. Yet, for other countries such as Singapore and Australia in which the fall period results are generally the same as the subperiod results, it would tend to suggest that the BDS test analysis was not influenced by regime changes.

In general, the BDS testing procedure results suggest that all of the national stock market indices, except Australia, may exhibit some form of nonlinearity. The FTSE, S&P 500, S&P 100, NASDAQ, DAX, and Nikkei indices have been tested for nonlinearity using the BDS test in prior research and the results are broadly consistent.\textsuperscript{15}

5.3. Close returns test results

As an alternative to the BDS test of nonlinearity, we may apply the close returns test as detailed in Section 4. Thus, all differences of the returns data for each stock index were computed according to $|x_t - x_{t+1}|$ where $t = 1 \ldots n$, $i = 1 \ldots n - 1$, and $n$ is the sample length. A summary plot of this information was constructed and Figs. 2 and 3 present the close returns data for Japan and Australia markets, respectively, which are representative of the type of plot observed for all 12 market indices tested. The test for the presence of chaotic behaviour relates to the identification of horizontal line segments. The close returns plots presented clearly do not possess such horizontal line segments such as those evident in Fig. 1 and this result was consistent for all countries tested and was not sensitive to the choice of threshold term.

While horizontal lines are not a feature of the close returns plot for the national stock market data considered in this paper, the data do exhibit some structure as the observations are clearly not random as would be the case if the data were IID. While the nature of the structure and the patterns evident in the data are a subject of ongoing research, it is known that if the structure were linear, the close returns plot would exhibit a clustering of points

\textsuperscript{15} See Abhyankar et al. (1995, 1997), Mayfield and Mizrach (1992), Sewell et al. (1996), and Willey (1992).
around the axis, which is clearly not the case in Figs. 2 and 3 or any of the other (unreported) close returns plots.\footnote{Diagonal segments are common to close returns plots (and nonzero base lines in the corresponding histogram) and are due to the result that in stationary data sets, upward trending segments of data are always followed by downward trending segments of data, which cause accidental close returns which can be cleaned up by embedding such as differential phase space embedding or an integral–differential filter. The close returns test is applied to a time series without an embedding and, if necessary, is embedded in Phase 2 where evidence of a strange attractor is found (Mindlin & Gilmore, 1992, p. 232).}

A formal quantitative test may be engaged, which consists of creating a histogram that summarises the number of close returns across each row in the close returns plot. If the data are IID, the histogram will be uniformly distributed around its mean value. As an illustration of this type of analysis, Fig. 4 presents the corresponding histogram generated from the close returns plot for the Japanese stock market data. To establish whether the number of close returns is uniformly distributed around the average number of close returns (presented as a horizontal line in Fig. 4), a chi-squared test may be estimated. The test results for all 12 stock market indices are presented in Table 3, and for the S&P 100, S&P 500, Nikkei Dow, Straits Times, and Swiss indexes as well as the Hang Seng (at the 10% level), the $\chi^2$ tests reject the null of IID. For the remaining six national stock markets, the histogram test fails to reject the null of IID data.

The more general box-plot test for nonlinearity, which divides the close returns plot into equal size boxes and compares the expected number of close returns to the actual, may also
be applied. The results of this procedure are presented in Table 4 and the \( P \) value associated with the \( \chi^2 \) test is less than .05 in every case tested. This is strong evidence in support of the alternative hypothesis of the presence of nonlinear behaviour and supports the casual observation of structure in the close returns plots.

In contrast to the BDS test, which provided no evidence of nonlinearity for the Australian market, and some evidence of nonlinearity for the US, UK, French, Canada, Germany, Hong Kong, Japan, Singapore, and Switzerland, the close returns test provides strong evidence of nonlinearity in each national stock market index although it is not chaotic. The Australian index provides a good example of the different results provided by these two testing methodologies. The BDS test failed to reject the null in each instance, yet the box-plot \( \chi^2 \) test was significant across all box sizes (the lowest ratio of \( \chi^2_i/\chi^2_c \) was 3.08 for the UK index with a box size of 20). The close returns plot for the Australian market presented in Fig. 3 would support this finding as the data plot appears to exhibit a nonrandom structure, which is to be expected where nonlinearity is present. Thus, the evidence obtained in this paper would tend to suggest that when considering testing for chaos and nonlinearity in small, noisy financial data sets, the close returns test is a useful tool that can compliment other nonlinear testing procedures, which were designed for large, clean data sets such as those found in the physical sciences.

As the close returns test for nonlinearity and chaos maintains the temporal ordering of the data, the \( \chi^2 \) test results are less susceptible to the presence of regime shifts compared to metric tests such as the BDS test. Nonetheless, it is possible that the detection of nonlinearity
may be due to the presence of nonstationarity. As such, subperiod analysis was undertaken as a means of testing whether the results vary with the sample. To this end, the box-plot $\chi^2$ test was applied to each half of the data sample. The results (not presented to conserve space) reveal that for both subperiods, the estimated $\chi^2$ coefficient was greater than the critical value for each box size in each index tested, which is entirely consistent with the full sample box-plot test results. The close returns plots also exhibited structure similar to the full sample box-

Table 3
Close returns histogram test results

<table>
<thead>
<tr>
<th>Index</th>
<th>$\chi^2$</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>509.22</td>
<td>.366</td>
</tr>
<tr>
<td>Canada</td>
<td>446.11</td>
<td>.956</td>
</tr>
<tr>
<td>France</td>
<td>418.39</td>
<td>.996</td>
</tr>
<tr>
<td>Germany</td>
<td>530.93</td>
<td>.156</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>549.97</td>
<td>.056</td>
</tr>
<tr>
<td>Japan</td>
<td>636.84</td>
<td>.000</td>
</tr>
<tr>
<td>Singapore</td>
<td>598.19</td>
<td>.001</td>
</tr>
<tr>
<td>Switzerland</td>
<td>598.43</td>
<td>.001</td>
</tr>
<tr>
<td>UK</td>
<td>493.48</td>
<td>.561</td>
</tr>
<tr>
<td>US S&amp;P 100</td>
<td>780.67</td>
<td>.000</td>
</tr>
<tr>
<td>US S&amp;P 500</td>
<td>728.81</td>
<td>.000</td>
</tr>
<tr>
<td>US NASDAC</td>
<td>528.93</td>
<td>.171</td>
</tr>
</tbody>
</table>

The following table presents the estimated $\chi^2$ test statistics and $P$ value for the test of the null hypothesis of independence applied to the histogram estimated from the close returns plot generated for each national stock market index over the period January 1990–December 1998.
Table 4
Close returns box plot test results

<table>
<thead>
<tr>
<th>Series</th>
<th>Box size</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>45</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>2600.13</td>
<td>2297.28</td>
<td>1848.40</td>
<td>1664.95</td>
<td>1360.16</td>
<td>1687.13</td>
<td>1382.89</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>2893.41</td>
<td>2394.86</td>
<td>1943.57</td>
<td>2169.55</td>
<td>1655.41</td>
<td>1514.84</td>
<td>1196.64</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>5453.96</td>
<td>5121.77</td>
<td>4506.21</td>
<td>3660.33</td>
<td>3513.23</td>
<td>3316.95</td>
<td>3563.76</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>11553.10</td>
<td>11236.40</td>
<td>10534.89</td>
<td>10402.05</td>
<td>9430.54</td>
<td>9875.89</td>
<td>10202.40</td>
<td></td>
</tr>
<tr>
<td>Hong Kong</td>
<td>3793.73</td>
<td>3416.55</td>
<td>2532.99</td>
<td>2596.61</td>
<td>2222.83</td>
<td>2238.73</td>
<td>2212.95</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>8282.14</td>
<td>8157.80</td>
<td>6666.19</td>
<td>6770.63</td>
<td>6047.35</td>
<td>5281.03</td>
<td>5833.09</td>
<td></td>
</tr>
<tr>
<td>Singapore</td>
<td>5916.24</td>
<td>5256.22</td>
<td>4753.64</td>
<td>3520.26</td>
<td>3549.39</td>
<td>3498.26</td>
<td>3683.61</td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>6684.76</td>
<td>5778.59</td>
<td>5166.66</td>
<td>5111.27</td>
<td>4631.92</td>
<td>4542.74</td>
<td>4602.76</td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>2108.35</td>
<td>1663.12</td>
<td>1368.16</td>
<td>1060.35</td>
<td>1228.18</td>
<td>1083.69</td>
<td>1025.56</td>
<td></td>
</tr>
<tr>
<td>US S&amp;P 100</td>
<td>3750.34</td>
<td>3610.00</td>
<td>2944.53</td>
<td>2704.06</td>
<td>2580.40</td>
<td>2871.79</td>
<td>2768.41</td>
<td></td>
</tr>
<tr>
<td>US S&amp;P 500</td>
<td>3441.84</td>
<td>3322.93</td>
<td>2763.85</td>
<td>2600.99</td>
<td>2381.63</td>
<td>2730.45</td>
<td>2480.86</td>
<td></td>
</tr>
<tr>
<td>US NASDAC</td>
<td>2683.67</td>
<td>2141.23</td>
<td>1683.92</td>
<td>1584.19</td>
<td>1148.28</td>
<td>1323.20</td>
<td>985.68</td>
<td></td>
</tr>
</tbody>
</table>

The following table presents the estimated $\chi^2$ test statistics and associated $P$ value (in parentheses) for the box plot test of the null hypothesis of independence applied to the close returns plot generated for each national stock market index over the period January 1990–December 1998.
plots. This result suggests that the nonlinearity detected by the close returns testing procedure is unlikely to be a reflection of any regime shift in the data. Compared to the BDS test results, which generated conflicting results such as those for Canada, the consistency of the close returns test results suggests that nonlinearity is present in all of the data sampled and the length and noise content of finance data may lead to inconsistent results in the context of BDS testing for nonlinearity.

6. Conclusion

Until recently, financial market researchers were ill equipped to detect the presence of chaos. The most commonly used nonlinear testing procedure was the BDS test, which is poorly suited for application to the small, noisy data sets common in finance. The introduction of the close returns test procedure for chaotic behaviour, however, has provided researchers with an exciting new tool for detecting chaos in finance data. This paper applied the BDS test procedure to 12 national stock market indices and the results suggest the presence of nonlinearities for all indices except the Australian market index. The close returns testing procedure was also applied to these data and the results furnished strong evidence of nonlinearity although it was not found to exhibit sensitive dependence on initial conditions, i.e., was not chaotic. In comparison to the BDS test results, the close returns test procedure provided much stronger evidence of nonlinearity in the data. The results provided by the BDS and close returns testing procedures indicate the presence of nonchaotic nonlinear behaviour in stock markets. Based on this evidence, future research may best be spent establishing on how best to explain and model such nonlinearities.

Acknowledgments

The author would like to thank Heather Mitchell, Claire Gilmore, and Robert Gilmore.

References


