Time series analysis of S&P 500 index: A horizontal visibility graph approach

Michail D. Vamvakaris [a], Athanasios A. Pantelous [a,b,*], Konstantin M. Zuev [c]

[a] *Department of Mathematical Sciences, School of Physical Sciences, University of Liverpool, Peach Street, Liverpool, L69 7ZL, UK*
[b] *Department of Econometrics and Business Statistics, Monash Business School, Monash University, Wellington Rd, Clayton, Victoria 3800, Australia*
[c] *Department of Computing and Mathematical Sciences, California Institute of Technology, 1200 E. California Blvd. Mail Code 305-16, Pasadena, CA 91125, USA*

HIGHLIGHTS

- The behavior in the S&P 500 price-index series is chaotic.
- The index series for the whole-time interval is an irreversible process.
- The behavior during a crisis occurs in the outer environment is a chaotic process.
- The impact of an external crisis affects the US market considerably.
- The US market never reaches equilibrium and remains in a disorder state constantly.

ARTICLE INFO

ABSTRACT

The behavior of stock prices has been thoroughly studied throughout the last century, and contradictory results have been reported in the corresponding literature. In this paper, a network theoretical approach is provided to investigate how crises affected the behavior of US stock prices. We analyze high frequency data from S&P500 via the Horizontal Visibility Graph method, and find that all major crises that took place worldwide in the last twenty years, affected significantly the behavior of the price-index. Nevertheless, we observe that each of those crises impacted the index in a different way and magnitude. Interestingly, our results suggest that the predictability of the price-index series increases during the periods of crises.

1. Introduction

For many decades now, there is an increasing interest both in academia and financial industry to what extent the stock's price history can assist to understand the behavior of its future prices. Particularly, throughout the previous century, various *chartist* theories were proposed assuming that the past behavior of a stock's price is rich of information concerning its future predictability. Thus, many researchers repeat the conjecture that "*patterns*" of past price behavior will tend to happen repeatedly in the future. On the other hand, the theory of random walks, developed by famous mathematicians, is used by

---

* Corresponding author at: Department of Mathematical Sciences, School of Physical Sciences, University of Liverpool, Peach Street, Liverpool, L69 7ZL, UK.

E-mail addresses: mvamvak@liverpool.ac.uk (M.D. Vamvakaris), a.pantelous@liverpool.ac.uk, Athanasios.Pantelous@monash.edu (A.A. Pantelous), kostia@caltech.edu (K.M. Zuev).

some economists to show that the stock's future price path is no more predictable than the path of a series of cumulated random numbers. In simple words, the past cannot be used to predict the future in any meaningful way.

In the direction of random walk, in 1900, the French mathematician Louis Bachelier [1] initiated a new chapter in the financial literature by proposing, however in a rather ambiguous way, that security prices appear to follow a stationary Gaussian random walk, and therefore price changes are uncorrelated. Many decades later, Bachelier's suggestion has been supported in a more explicit way by Osborn's work [2], where he show that the change of logarithm of prices has the same properties as a particle in the Brownian motion.

The Gaussian hypothesis was seriously questioned by Mandelbrot [3,4] who observed that the implication of the *leptokurtosis* (i.e., its "excessively peaked" character) was neglected by the previous researchers, contrarily on what someone could observe in the empirical distribution of price changes. In another words, his findings suggest that the distribution of the logarithm of price changes is *not* Gaussian, but rather *stable* Paretian since outliers are actually more common than there would expect to be under the Gaussian distribution. Additionally, Mandelbrot [4] provided evidences that stock price series are not stationary, and there is a weak (linear or nonlinear) dependence in their changes. Although Fama [5,6] concluded that stock price changes do follow a stable Paretian distribution, he insisted that daily price increments are random, independent and not correlated.

Many years later, Lo and MacKinlay [7] tested the random walk hypothesis for weekly price changes data on an equal- and a value-weighted market indexes. For the entire period and for all subperiods, the rejection of the random walk hypothesis did not support a mean-reverting model of asset prices. Thus, they alleged that the presence of heteroscedasticity was not able to explain completely the rejection of the random walk hypothesis. What is more, they found a positive serial correlation in contrast with Fama and French [5], who observed a weakly negative autocorrelation in weekly stock price changes. The negative autocorrelation was appeared due to the slowly mean-reverting behavior of the stock price series while they also reported that autocorrelation varied over time. Since then, a voluminous research has been conducted to confirm or reject the validity of the random walk hypothesis based on the analysis of empirical data.

This paper is inspired by results obtained in Cutler et al. [8], who demonstrated that about one third of the stock price changes in S&P500 can be attributed to the arrival of macroeconomic and political news. Furthermore, despite the passage of time and the massive improvement in technology, Cutler's et al. [8] conclusion remained unchanged even though the size of unexplained market movements has grown incredibly, as Cornell [9] demonstrated recently. In this direction, we study how selective events in the course of time have affected the behavior of S&P500 price-index using high frequency data for the time period between May 1996 and March 2016. Instead of just using tests to reject or confirm the random walk hypothesis, we take advantage of the recently introduced Horizontal Visibility Graph method [10,11] to identify the characteristics of the generating process of the index prices.

The main purpose of our study is to investigate whether crises have affected the behavior of the index, and if so, to examine whether different types of crises had the same impact on the index. In this regard, using 5-min data and applying the rolling window method with overlap, we find that all major crises that took place the last twenty years have affected considerably the behavior of the index. Moreover, our results suggest that there exist certain patterns in the way that the index reacts to specific crisis events.

The remainder of the paper is structured as follows. In Section 2, we provide the literature review of the Horizontal Visibility Graph method which is the cornerstone of our study. In Section 3, the presentation of the data is provided. A detailed analysis of the adapted methodology is given in Section 4. Empirical results and their analysis for the whole time period, and the irreversibility of the S&P 500 price-index are presented in Section 5. Section 6 concludes the whole discussion.

2. Network theoretical approach: Horizontal visibility graph

Network modeling is a novel and rapidly developing area of research in social sciences, economics and finance. Mantegna and his co-authors [12–17] were the first who used techniques and tools from network theory to uncover the true nature of the underlying structure of financial markets. Many more researchers have followed, see [18–25] among numerous others.
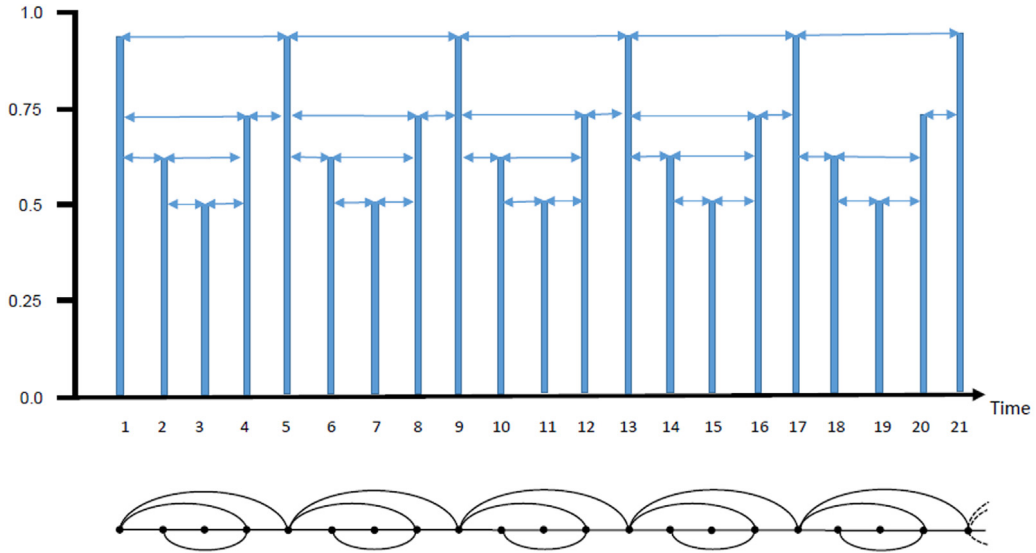
In our paper, the *Horizontal Visibility Graph* (HVG) which was recently introduced by Luque et al. [10] is considered. The HVG algorithm belongs to the family of *visibility algorithms* (VG) proposed by Lacasa et al. [26], and uses the concept of visibility [27] to map time series into graphs according to a specific geometric criterion. In [26], it was shown that time series structure is inherited in the associated graph, such as periodicity, randomness, etc. Thus equivalently, one can extract useful information regarding the properties of the original series based on the analysis implemented in the associated graph.

The geometric criterion proposed for the HVG method is basically rather simple, see [10]. Let $\{X_i\}_{i=1,2,...,n}$ be a time series with n observations. Each point in the series is transformed into a node in the associated graph, and two arbitrary nodes i and j have "visibility" to each other, and hence they are adjacent in the associated graph, if:

$$x_i, x_j > x_n \text{ for every } n \text{ such that } i < n < j.$$

For illustrative purposes, a graphical representation of the HVG method is shown in Fig. 1. In the upper part, there is a bar-chart of a periodic time series while in the lower part there is a depiction of the associated graph after the HVG method was applied to the original data. Based on Gutin et al. [28], we can emphasize two of the most significant properties of HVG:

1. The associated graph is always connected, since all data points in the series have visibility to the nearest neighbors (i.e., the previous and the following one).

Fig. 1. Graphical illustration of the Horizontal Visibility Graph algorithm. On the top we plot a periodic time series while at the bottom there is a graphical representation of the associated network.

2. The mapping process is invariant under any affine transformation of the original time series, i.e., we will end up with the same graph regardless of any axis rescaling of the initial series.

In the next subsections, the importance of the properties mentioned above for the analysis of financial data is explained.

2.1. Description of randomness and chaos

In [10,29], it is shown that the associated graph of any uncorrelated stochastic (random) series, which is independent of the shape of the probability density function $f(x)$ (i.e., distribution-free), has the degree distribution of the form:

$$P(\kappa) = \frac{1}{3}\left(\frac{2}{3}\right)^{\kappa-2},\qquad(1)$$

where $\kappa \geq 2$ for series of infinite length. Easily, Eq. (1) can be rewritten as

$$P(\kappa) \sim \exp^{-\gamma\kappa},\qquad(2)$$

where in the case of uncorrelated stochastic series, $\gamma = \ln(3/2)$. For *correlated* stochastic process, the exponent $\gamma > \ln(3/2)$; higher values of γ correspond to stronger *autocorrelation* in the series. On the other hand, for series that derives from *chaotic* processes, the exponent $\gamma < \ln(3/2)$, and higher values of γ correspond to higher dimensions of the attractor.

In Fig. 2, a graphical illustration of the degree distribution of the associated HVG for three different series, random, correlated stochastic and chaotic, is provided. Based on [29], the exponent γ can be calculated by fitting an exponential function to the tail of degree distribution of the associated graph. In the present paper, the exponent γ is calculated by fitting a linear function to the tail of the degree distribution in a semi-log scale. Numerical experiments in [10,29] show that the HVG method is a reliable tool to distinct both noisy and high dimensional chaos from randomness (stochasticity).

2.2. Quantifying the degree of irreversibility of time series

A time series $S = \{x_1, x_2, \ldots, x_n\}$ is called statistically time-*reversible* if $S^- = \{x_{(-1)}, x_{(-2)}, \ldots, x_{(-n)}\}$ has the same joint distribution with S [30]. This merely implies that both S and S^- series are statistically equally probable, and consequently, any independent and identical distributed random variable is a reversible sequence. Using another definition, we can conclude that "*the probabilistic properties of reversible series are invariant with respect to time reversal*" [31], and that "*reversibility is shown to imply stationarity*" [32]. Based on the aforementioned definitions, nonstationary time series are always *irreversible*, and thus it is only meaningful to study time reversibility in stationary series. Stationary processes, such as the linear Gaussian random process, that exist at thermal equilibrium are reversible [33]. On contrary, stationary processes which contain underlying nonlinearities such as dissipative chaos, and other processes that exist out of thermal equilibrium are irreversible processes. In this regard, many researchers also claimed that time irreversibility is associated with the presence of nonlinearity, and in particular, Cox [34] claimed that "*irreversibility is the symptom of nonlinearity*". The HVG method
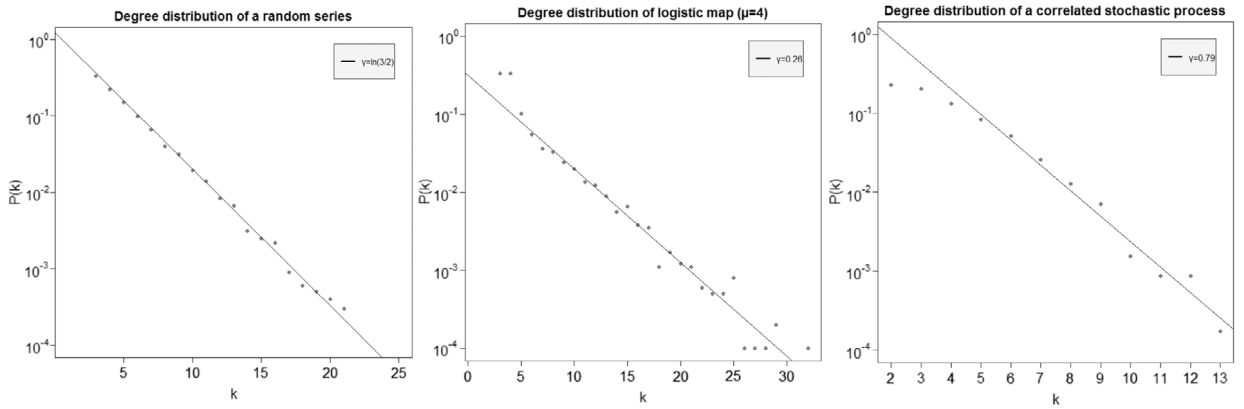
Fig. 2. Degree distribution of (from left to right) (a) random process, (b) chaotic process, (c) correlated stochastic process. The length of each series is $2 \cdot 10^4$.

can not only successfully discriminate between reversible and irreversible stationary time series, but it can also quantify the degree of time irreversibility for nonstationary time series [33]. Since the definition of a reversible series requires stationarity, by implementing the new concept of HVG reversibility here, we are also able to quantify the degree of time asymmetries in the underlying dynamics of a nonstationary time series. Based on this concept, a process is characterized as time reversible when both direction of time evolution seems "*right*" for a particular series.

Following [35], a dynamical process $\{X_t\}$ is said to be HVG stationary *if and only if* the topological properties of the HVG associated to a sample time series of size n extracted from $\{X_t\}$ are asymptotically (for large n) invariant under time shift (in the statistical sense). More specifically, let $S = \{x_1, x_2, \ldots, x_n\}$ and $S^\tau = \{x_{1+\tau}, x_{2+\tau}, \ldots, x_{n+\tau}\}$ be two sample series extracted from $\{X_t\}$. The degree distributions of HVG associated to the sample series, S and S^τ, are asymptotically (for large n) identical for HVG stationary processes.

To be able to discriminate between reversible and irreversible time series, one has to separate the degree of each node into κ_{in} and κ_{out} such that $\kappa(i) = \kappa_{in}(i) + \kappa_{out}(i)$, where $\kappa_{in}(i)$ is the number of nodes from the past which have visibility on a node i, and $\kappa_{out}(i)$ is the number of future nodes where a node i has visibility based on the HVG criterion. Hence, a process $\{X_t\}_{t=1,2,\ldots,n}$ is assumed to be HVG reversible *if and only if* the in-degree and the out-degree distributions of the associated HVG are *asymptotically identical* for large values of n. For series of finite size, we can quantify the degree of irreversibility of a process by calculating the divergence between the in- and out-degree distributions of the associate HVG. The higher the distance between the two distributions is, the more irreversible a process is.

In our study, instead of employing the Kullback–Leibler divergence for calculating the divergence between the in- and out-degree distribution of the HVG as originally proposed in [33], we recruit the Hellinger distance [36], which does not introduce any bias for time series of finite length [37]. The Hellinger distance between two discrete probability distributions, $P(p_1, p_2, \ldots, p_n)$ and $Q(q_1, q_1, \ldots, q_n)$, is given by:

$$H(P, Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^{n} (\sqrt{p_i} - \sqrt{q_i})^2}. \tag{3}$$

The value of Hellinger distance should be zero for HVG reversible process of infinite size. However, given the fact that all real life time series are finite, it is possible to face the situation where a reversible series has $H \neq 0$ yet, as the size of the series increases $H \to 0$ asymptotically. If a series is an irreversible one, then, as the size of the series increases, $H \to h$ asymptotically, where $h > 0$. In Fig. 3, we plot the in- and out-degree distribution of a reversible and an irreversible process. In the first case, although the two distributions are very close to each other, they are not identical and this can be attributed to the finite length of the series. For the irreversible process, the divergence between the distributions is obvious.

3. Data

Our base dataset consists of tick data from Thompson Reuters Tick History (TRTH) for the S&P500 price-index that starts on the 3rd May 1996 to 3rd March 2016. Since higher frequency inevitably leads to a larger microstructure noise, as it has been suggested by Hansen and Lunde [38], we initially aggregate the data and construct a dataset that contains 440,000 observations of 5-min data points, see also arguments in [39]. In Fig. 4, we plot the daily closing prices of the index for the corresponding time interval. It is very common in economics and finance to exploit the properties of logarithms, and instead of analyzing the original dataset, to work with the logarithm of the data. As it is already mentioned in Section 2, HVG method is invariant under axis transformation, and as a direct consequence, both time series, i.e., with original prices and with log-prices, generate the same HVG graph.
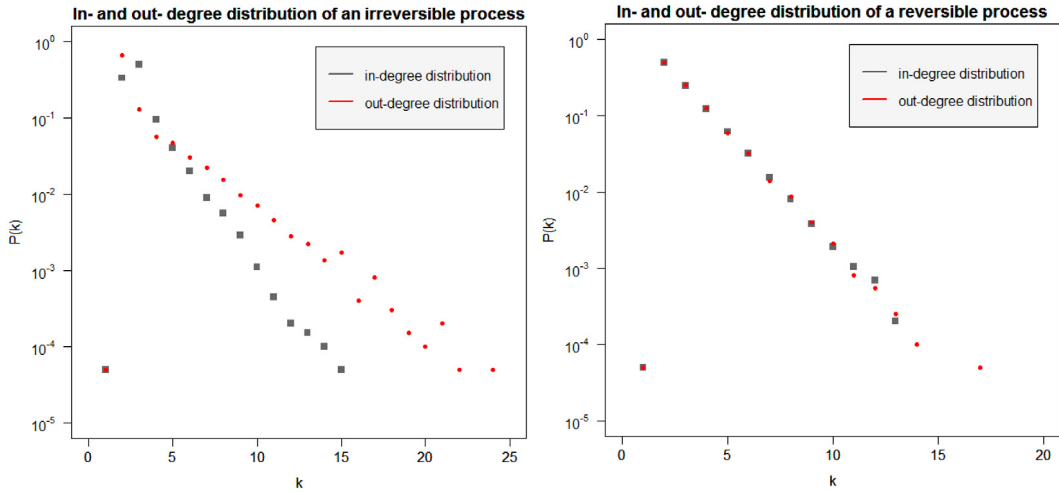
Fig. 3. The in- and out-degree distributions of an irreversible process (left) and a reversible process (right). The reversible series is a sequence of a random numbers drawn from the uniform distribution on [0, 1] while the irreversible series is a sequence of numbers generated by the logistic equation with $\mu = 4$. The length of both series is $N = 2 \cdot 10^4$.
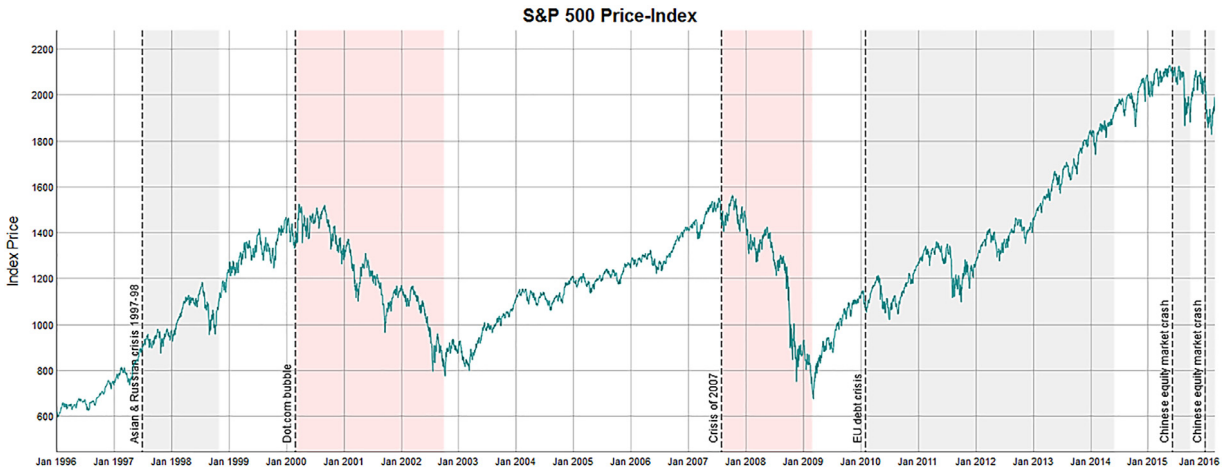

Fig. 4. S&P 500 price-index series from 3rd May 1996 to 3rd March 2016. The red shadowed areas denote periods of sovereign crisis in USA while the gray shadowed areas denote crisis that took place abroad. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Methodology

In the present paper, the question on how particular crises occurred during the last 20 years affected the behavior of the S&P500 is considered. In this regard, we split our initial dataset into shorter periods of study to identify possible changes in the US market behavior. Moreover, in order to achieve continuity in our results, the method of rolling window with overlap is applied. Particularly, the length of each window is 5800 observations (or 71 trading days), and the overlap between two consecutive windows is 3900 observations (or 48 trading days). In total, we create 229 subperiods of interest, and each and every subperiod is transformed into a graph by using the HVG method. For every window, the exponent γ of the degree distribution, Eq. (2) is calculated as well as the Hellinger distance between the in- and out-degree distribution of the associate network, Eq. (3). Throughout this process, the properties of the series can be identified, i.e., if the series is correlated (or uncorrelated) stochastic or chaotic, and the degree of irreversibility of each sub-period can be quantified as well. Finally, whether the behavior of the sub-series has changed during the periods of crisis is examined, and whether the reaction of

Table 1
Subperiods with the six major crises considered.

Subperiods	Time periods
Subperiod 1 (Asian & Russian crisis):	July 1997–December 1998
Subperiod 2 (Dot.com bubble collapse):	March 2000–November 2002
Subperiod 3 (USA subprime crisis):	September 2007–January 2009
Subperiod 4 (European debt crisis):	October 2009–February 2014
Subperiod 5 (Chinese stock market crash, Face I)	June 2015–September 2015
Subperiod 6 (Chinese stock market crash, Face II)	January 2016–March 2016
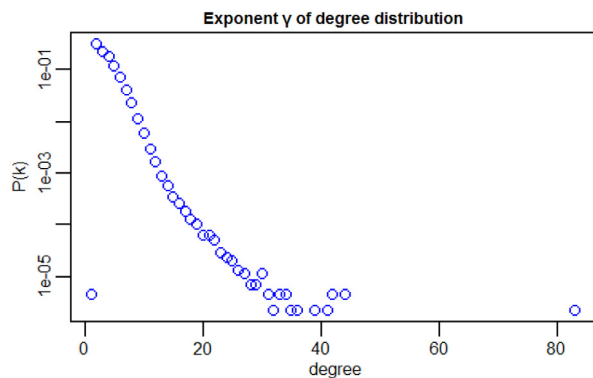
Fig. 5. Semi-log plot of the degree distribution of the HVG of the S&P500 price-index from 1996 to 2016.

the market is the same for all crisis events is tested. For the better understanding of the empirical part, in Table 1, six major crises occurred throughout the time-period of observations are reported.[1]

5. Empirical analysis

In this section, the empirical analysis for S&P500 case implementing the HVG method is conducted. The section is divided into two subsections. In the first part, the market behavior is analyzed for the whole depicted period as well for sub-periods 1 to 6 paying attention particularly on how crises affected the S&P 500 price-index. In the second part, whether S&P 500 price-index is a time reversible series, and how crises affected the degree of its irreversibility in the course of time are investigated.

### 5.1. S&P 500 index series all over the period 1996–2016

After using the HVG method, i.e., mapping the observations of S&P 500 price-index time series in a graph, we are in a position to extract the degree distribution of the associated graph. Then, the value of the exponent γ, see Eq. (2), is calculated, which is 0.026 for the whole time interval starting from 3rd May 1996 to 3rd March 2016. In Fig. 5, the degree distribution of the associate graph of the index series is plotted.[2] The exponent γ equals the slope of the tail of the degree distribution in a semi-log plot.

The fact that the value of exponent γ is less than $\ln(3/2)$ explains that the S&P500 price-index series appears to exhibit a chaotic behavior, and more specifically, it has similar properties with a low dimensional chaotic process. Voluminous literature exists providing some evidence that stock price series behave similar to a chaotic process, and the existence of long term dependences can be confirmed within stock price series [40–43]. However, the existing statistical tests cannot accurately infer whether a series is indeed chaotic or random. Our finding does not contradict, but rather verifies and reasonably explains what other researchers have proposed using other methods.

Our results reveal that a deterministic nonlinear structure underlies the S&P500 high-frequency data. This outcome explains why many researchers traced a nonlinear dependence in stock price series, yet they were uncertain for the source of this nonlinear structure [44]. The existence of chaos might convincingly explain the nonlinear dependencies derived in financial time series. However, there are contradicting studies as well, see [45–47].

---

[1] Since by the end of February 2014 four out of five European countries that had asked for a bailout programme accessed again the private debt market, we assume that the European debt crisis (see Subperiod 4) is over.

[2] As a minor remark we should mention here that there is a node with extreme degree of 1750, which has not been included in Fig. 5 for visualization arguments. However, the analysis conducted for the exponent γ considers also that extreme value.
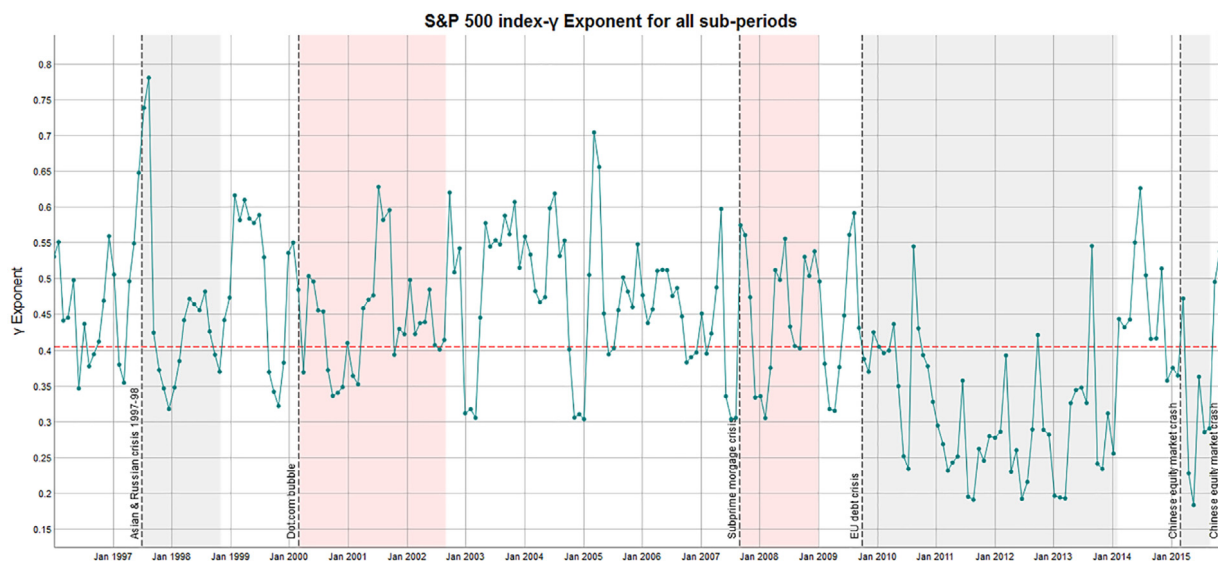
Fig. 6. Values of exponent γ for all windows for the S&P 500 price-index. In the vertical axis it is plotted values of γ exponent and on the horizontal line we plot the first date of each window. The red line represents the value $\gamma = \ln(3/2)$ which correspond to an uncorrelated random process. Red shadowed regions represent periods of market crashes following a US market bubble, and gray shadowed regions represents periods of crisis that took place in the economic/political reality of USA or abroad. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Number (ratio) of windows where exponent γ lies in the corresponding band.

	Chaos	Uncorrelated stochastic	Correlated stochastic
No of windows	75 (32.75%)	2 (0.87%)	152 (66.38%)

The most important implication of the chaotic behavior is that stock prices appear to be predictable to some extent. However, the horizon where we can successfully predict the future trajectories can be better determined by calculating the Lyapunov exponents. Even though a chaotic process is indistinguishable from the unpredictability, due to the sensitivity issues in the initial conditions, a chaotic system is predictable for a limited time span and then its behavior is seemingly random. Despite its phenomenal randomness, any chaotic system does contain a deterministic structure. In addition, in classical thermodynamics chaos exists in dynamical systems that operate out of equilibrium and are characterized by entropy production. Pro rata, as the equity market behaves similar to chaos, it is a dynamic system that continuously operates out of equilibrium. Thus, S&P 500 price-index can be seen as a non-isolated system where the outer environment interrupts its way to equilibrium by introducing external turbulence in the system and causing entropy production.

In Fig. 6, we observe the values of exponent γ for all distinct time-windows. At this point, we should mention that each *dot* represents a different window, and dots are located on the *x*-axis based on the date of the first observation in each window.[3] The *red line* denotes the point where γ exponent equals $\ln(3/2)$, which corresponds to an uncorrelated stochastic process. *Red shadowed areas* indicate periods of crises which have been endogenously created by the market forces, i.e., large drop in stock prices following a price bubble. On the other side, *gray shadowed areas* represent periods of crises when the crisis occurred either in the US economic/political reality or abroad. Consequently, Dot.com bubble collapse (2000) and US subprime crisis (2007) are denoted in the graph with red color while all the other crisis events are denoted by gray color.

With a glance on Fig. 6, it is easily discernible that in the gray shadowed areas there are valleys of the γ exponent in the chaotic band, a fact which is not true in the red shadowed areas. In Table 2, a summary of the market behavior is reported during the periods considered. However, since our major endeavor is to study the behavior of the market during periods of crisis, in Table 3, we also report the ratio of windows that exhibit a particular behavior and coincide with a crisis event.

What we can infer from Table 3 is that mainly chaotic behavior emerges when a crisis is ongoing, while correlated stochastic behavior is more common when there is no an ongoing crisis. About 83.18% (or 94 out of 113) of windows that coincide with tranquil periods appear to behave similar to a correlated stochastic process.[4] When a crisis is ongoing, almost

---

[3] For example, the first window covers the period from 3rd May to 8th August 1996 and the corresponding dot is located on the X-axis at the point 3rd May 1996.

[4] We do not make any comment regarding the uncorrelated stochastic behavior since it does not seem to be a repeated behavior that deserves a particular analysis.

Table 3
Number (ratio) of windows where exponent γ lies in the corresponding band during periods of crisis and during tranquil periods.

	Chaos	Uncorrelated stochastic	Correlated stochastic	Total
Crisis Event	57 (76%)	1 (50%)	58 (38.15%)	116
Tranquil period	18 (24%)	1 (50%)	94 (61.85%)	113
Total	75	2	152	229

Table 4
Number (ratio) of windows where exponent γ lies in a particular band and coincide with a particular type of crisis event, i.e., crises endogenously created by the market or crises in the outer environment of the market.

	Chaos	Uncorrelated stochastic	Correlated stochastic
Endogenous crises	5 (8.8%)	1 (100%)	40 (68.96%)
Exogenous crises	52 (91.2%)	0 (0%)	18 (31.04%)

50% of the periods (windows) appear to behave similar to chaos and another 50% similar to a correlated stochastic process. An important aspect would be to investigate if there is any pattern in the behavior of the market during different types of crisis.

In Table 4, the set "crisis event" is partitioned into two non-overlapping sets; one group which contains crisis events that are created *endogenously* by the market forces, and another group with all the other crises that took place in the outer environment of the US equity market, i.e., crisis that raised either in the US economic and political reality or abroad.

Results in Table 4 indicate that chaotic behavior rises when a crisis occurs in the outer environment of the market, while the properties of the S&P 500 price-index are similar to a correlated stochastic process, when the crisis has been created endogenously by the market dynamics. In total, 86.95% (or 40 out of 46) of the windows that coincide with an internally created crisis by the market forces behave similar to a correlated stochastic process, while 91.2% (or 52 out of 57) of windows that coincide with crises happened in the US economic and political reality or abroad behave similar to a chaotic process.

Those findings reveal some very interesting aspects of the US market behavior. Bearing in mind that almost 83% of the windows that coincide with tranquil periods behave similar to a correlated stochastic process, and that around 87% of the windows that coincide with crisis events created internally by markets' dynamics, then we have some evidence to infer that these types of events do *not* alternate the behavior of the market at all. Additionally, it appears the behavior of the market does not alternate when a bubble in stock price does inflate and deflate. Consequently, US market reacts in most cases, when an event happens in its outer environment, and exhibits a chaotic behavior. Consequently, as long as the existence of chaos is related to entropy production, and entropy is a measure of disorder, we can infer that events that take place in the outer environment of the market drive the system out of equilibrium causing an increasing disorder in the system. For instance, let us have a close look in an event that happened at the end of 2003 and beginning of 2004. As we can observe in Fig. 6, there is a rapid change in the behavior of the market similar to that of the Asian crisis in 1997, yet these windows are not shadowed by any color. At that time the Federal Reserve rose the interest rates, and at the same time oil reached its maximum price for the last 20 years. Although, there is no assurance for which of these events caused an alternation in market behavior or affected the most, there is a strong indication to conclude that chaotic behavior appears when an important event takes place in the outer market environment.

Before proceeding to the second part of this section, the analysis conducted so far reveals that the US market reacts only in the case that a crisis is ongoing either in the US political and economic reality or abroad, while equity market crises that follow a rapid stock price expansion do not alter the behavior of the market.

5.2. Irreversibility of S&P 500 price-index

In the previous subsection, the behavior of S&P500 price-index is studied over the sample period, and how particular events affect the index is considered by calculating the γ exponent of Eq. (2). In this section, another property of the index series will be explored, and thus, how the time irreversibility of the price-index series varies over the course of time is examined. As discussed in Section 2, this can be accomplished by calculating the Hellinger distance, H, between the in- and out-degree of the associate graph.

In Fig. 7, the in- and out-degree distributions of the directed HVG are plotted. What it is observed is that the two distributions do not coincide, and thus, we cannot allege that the series is a time reversible one. The calculated Hellinger distance between the two distributions is 0.0136, which is larger than zero verifying that the series is irreversible. We emphasize that given the large size of this series, the expected value of the Hellinger distance for a reversible series should be approximately zero. This result verifies what we derived in Section 5.1, i.e., that the market behaves as a chaotic process, and it is true that any chaotic process is an irreversible one.

However, for reversible series of finite size, Hellinger distance is not exactly zero yet the value of H approaches asymptotically zero as the length of a series tends to infinity. The challenge occurred at this point is that, to the best of
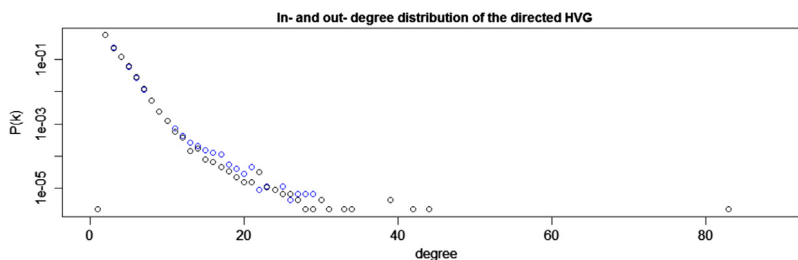
Fig. 7. The in-degree and out-degree distributions of the associated directed HVG of the S&P price-index from 1996 to 2016. Blue cycles represent the in-degree distribution and black cycles the out-degree distribution.
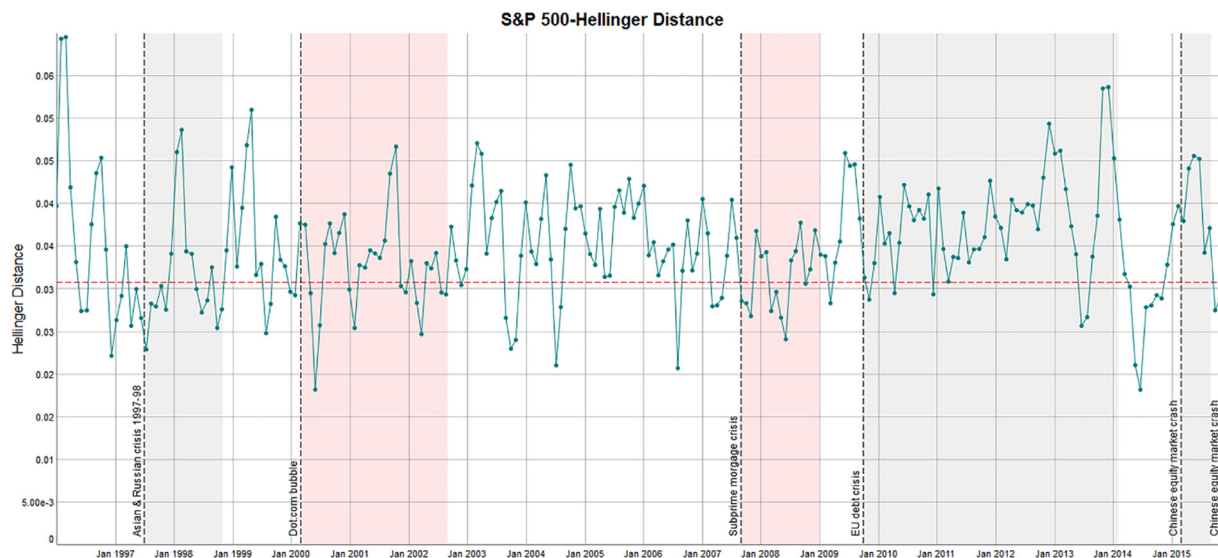
Fig. 8. Values of Hellinger distance for all windows of the S&P 500 price-index. In the vertical axis, the Hellinger distance values are plotted, and on the horizontal line we plot the first date of each window. Red shadowed regions represent periods of market crashes following a market bubble in USA and gray shadowed regions represents periods of crisis that took place in the economic/political reality of USA or abroad. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

our knowledge, there is no benchmark value available for H that separates reversible and irreversible series of finite size. Nevertheless, after having performed the Augmented Dickey–Fuller root tests [48], we found that only 18 out of 229 subseries are stationary with confidence 99%. This implies that we do not actually examine whether subseries is time reversible, but rather, how irreversible it is and how crisis affects time asymmetries. For clarity, we mention that a nonstationary process is always time irreversible.

In Fig. 8, we plot the values of the Hellinger distance for all windows. We observe that none of the values is close to zero which verifies that all subseries are irreversible. This finding is in line with [26,28] where it is stated that financial series are in principal irreversible. After all, it would be unreasonable to find reversibility in our results given that "*time irreversibility is the rule rather the exception when it comes to nonlinearity*" [29], and we just show that S&P 500 index behaves similar to a chaotic process. Interestingly, the irreversibility of the index varies considerably over time. Within the red shadowed areas, we observe that H fluctuates at a lower level in comparison with the white and gray regions. Exceptions are two windows at the end of 2001 where H peaks. These windows coincide with the 9/11 terrorist attacks, a major event in the recent US history that affected considerably the properties of the S&P 500 index series. On the other hand, within the gray shadowed areas, H fluctuates at higher level in comparison with the red shadowed. This is especially visibly during the European debt crisis and Chinese equity market crash. Tumminello et al. [14] found a link between the degree of time irreversibility of a process, and its amount of entropy production. The more irreversible a process is, the higher the production of entropy of the system is. Thus, we can infer that gray shadowed regions are periods when the market is more disordered in comparison to tranquil periods and periods of endogenously created by the markets' dynamics crisis.

What is more, increasing irreversibility denotes also increasing time asymmetry. Two of the most important asymmetries that underlie financial transactions are asymmetry in investors believes and information. Even though we cannot claim which kind of asymmetry is affected more by the existence of a crisis, we can allege that these types of crises increase asymmetries within the market.

Finally, the fact that red shadowed areas have a lower degree of time irreversibility denote less disorder in the market throughout the corresponding time interval. This might be as a contradiction with what we proposed in Section 5.1 observing that the behavior of the index remains the same during tranquil periods and periods where an endogenously created by the markets' dynamics crisis takes place. However, economists characterize these types of crises as a correction movement on behalf of markets' participants after a prolonged inefficiency of the market. Thus, this behavior is not unexpected.

6. Conclusions

In this paper, we recruit the recently introduced HVG method to analyze high frequency (5-min) US stock market data, and examine the behavior of the S&P 500 price-index series from May 1996 to March 2016, and investigate how different kinds of crises affect the behavior of the market.

Our findings provide strong evidence that behavior in the S&P 500 price-index series is chaotic, which means that the stock series has an underlying deterministic structure, yet a nonlinear one, and thus it is unpredictable. The fact that the index series for the whole time interval is an irreversible process provides even more support to our argument.

Since the main goal of our approach is to analyze the behavior of the index series throughout periods of crises, we decompose the initial series into six subperiods, and the behavior of each particular subperiod is examined. For the sake of continuity, in our results, the rolling window with overlap method is implemented where the length of each window is 71 trading days, and the overlap between two consecutive windows is 48 trading days.

We found that the behavior of the market during tranquil periods as well as during periods of endogenously created by the market dynamics crises, i.e., large drop in stock price following a bubble in the stock prices, is similar to a correlated stochastic process with strong autocorrelation among prices. On the other hand, when a crisis happens in the outer environment of the market i.e., either in the US economic and political reality or abroad, then, the stock price series behaves similar to a chaotic process. In other words, when a bubble in the stock prices burst it does not affect particularly the behavior of the index, while changes in the state of the economy or politics of the country or abroad drive the market out of equilibrium and increases the disorder in the system.

Similar conclusions can be derived after examining how irreversibility of series evolve over the course of time. The impact of an external crisis affects the market considerably, and the degree of irreversibility reaches its maximum values which means that the disorder in the market is high since a link between irreversibility and entropy production exists. On contrary, when a market experiencing a stock price bubble burst, the irreversibility fluctuates at low levels indicating the disorder in the system is generally much less than any other period.

Finally, the most important conclusion is that since none of the subseries seems to be reversible, and consequently the market increases its entropy, the market never reaches equilibrium and remain in a disorder state constantly.

References

[1] L. Bachelier, Théorie de la spéculation, Gauthier-Villars, 1900.
[2] M.F.M. Osborne, Brownian motion in the stock market, Oper. Res. 7 (2) (1959) 145–173.
[3] B. Mandelbrot, The variation of some other speculative prices, J. Bus. 40 (4) (1967) 393–413.
[4] B. Mandelbrot, The variation of certain speculative prices, in: Fractals and Scaling in Finance, Springer, 1997, pp. 371–418.
[5] E.F. Fama, K.R. French, Permanent and temporary components of stock prices, J. Polit. Econ. 96 (2) (1988) 246–273.
[6] E.F. Fama, Random walks in stock market prices, Financ. Anal. J. 51 (1) (1995) 75–80.
[7] A.W. Lo, A.C. MacKinlay, Stock market prices do not follow random walks: Evidence from a simple specification test, Rev. Financ. Stud. 1 (1) (1988) 41–66.
[8] D.M. Cutler, J.M. Poterba, L.H. Summers, What moves stock prices? J. Portfolio Manage. 15 (3) (1989) 4–12.
[9] B. Cornell, What moves stock prices: Another look, J. Portfolio Manage. 39 (3) (2013) 32.
[10] B. Luque, L. Lacasa, F. Ballesteros, J. Luque, Horizontal visibility graphs: Exact results for random time series, Phys. Rev. E 80 (4) (2009) 046103.
[11] M. Vamvakaris, A.A. Pantelous, K. Zuev, Investors' behavior on S&P 500 index during periods of market crashes: A visibility graph approach, in: F. Economou, K. Gavriilidis, G.N. Gregoriou, V. Kallinterakis (Eds.), Handbook of Investors' Behavior During Financial Crises, Academic Press, 2017, pp. 401–417.
[12] R.N. Mantegna, Hierarchical structure in financial markets, Eur. Phys. J. B 11 (1) (1999) 193–197.
[13] G. Bonanno, G. Caldarelli, F. Lillo, S. Miccichè, N. Vandewalle, R.N. Mantegna, Networks of equities in financial markets, Eur. Phys. J. B 38 (2) (2004) 363–371.
[14] M. Tumminello, T. Aste, T. Di Matteo, R.N. Mantegna, A tool for filtering information in complex systems, Proc. Natl. Acad. Sci. USA 102 (30) (2005) 10421–10426.

[15] F. Lillo, S. Miccichè, M. Tumminello, J. Piilo, R.N. Mantegna, How news affects the trading behaviour of different categories of investors in a financial market, Quant. Financ. 15 (2) (2015) 213–229.
[16] C. Curme, M. Tumminello, R.N. Mantegna, H.E. Stanley, D.Y. Kenett, Emergence of statistically validated financial intraday lead–lag relationships, Quant. Financ. 15 (8) (2015) 1375–1386.
[17] G. Iori, R.N. Mantegna, L. Marotta, S. Micciche, J. Porter, M. Tumminello, Networked relationships in the e-mid interbank market: A trading model with memory, J. Econom. Dynam. Control 50 (2015) 98–116.
[18] O.V. Precup, G. Iori, A comparison of high-frequency cross-correlation measures, Physica A 344 (1) (2004) 252–256.
[19] T. Aste, T. Di Matteo, Dynamical networks from correlations, Physica A 370 (1) (2006) 156–161.
[20] D. Garlaschelli, T. Di Matteo, T. Aste, G. Caldarelli, M.I. Loffredo, Interplay between topology and dynamics in the world trade web, Eur. Phys. J. B 57 (2) (2007) 159–164.
[21] G. Iori, G. De Masi, O.V. Precup, G. Gabbi, G. Caldarelli, A network analysis of the Italian overnight money market, J. Econom. Dynam. Control 32 (1) (2008) 259–278.
[22] F. Pozzi, T. Di Matteo, T. Aste, Spread of risk across financial markets: better to invest in the peripheries, Sci. Rep. 3 (2013) 1665.
[23] J. Birch, A.A. Pantelous, K. Zuev, The maximum number of 3-and 4-cliques within a planar maximally filtered graph, Physica A 417 (2015) 221–229.
[24] J. Birch, A.A. Pantelous, K. Soramäki, Analysis of correlation based networks representing dax 30 stock price returns, Comput. Econ. 47 (4) (2016) 501–525.
[25] A. Temizsoy, G. Iori, G. Montes-Rojas, Network centrality and funding rates in the e-mid interbank market, J. Financ. Stab. (2017) (in press).
[26] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J.C. Nuno, From time series to complex networks: The visibility graph, Proc. Natl. Acad. Sci. 105 (13) (2008) 4972–4975.
[27] M. De Berg, O. Cheong, M. Van Kreveld, M. Overmars, Computational Geometry: Introduction, Springer, 2008.
[28] G. Gutin, T. Mansour, S. Severini, A characterization of horizontal visibility graphs and combinatorics on words, Physica A 390 (12) (2011) 2421–2428.
[29] L. Lacasa, R. Toral, Description of stochastic and chaotic series using visibility graphs, Phys. Rev. E 82 (3) (2010) 036120.
[30] G. Weiss, Time-reversibility of linear stochastic processes, J. Appl. Probab. 12 (4) (1975) 831–836.
[31] C. Diks, J.C. Van Houwelingen, F. Takens, J. DeGoede, Reversibility as a criterion for discriminating time series, Phys. Lett. A 201 (2) (1995) 221–228.
[32] A.J. Lawrance, Directionality and reversibility in time, Internat. Statist. Rev. 59 (1) (1991) 67–79.
[33] L. Lacasa, A. Nunez, É. Roldán, J.M.R. Parrondo, B. Luque, Time series irreversibility: a visibility graph approach, Eur. Phys. J. B 85 (6) (2012) 1–11.
[34] D.R. Cox, Statistical analysis of time series: some recent developments, Scand. J. Stat. (1981) 93–115.
[35] L. Lacasa, R. Flanagan, Time reversibility from visibility graphs of nonstationary processes, Phys. Rev. E 92 (2) (2015) 022817.
[36] E. Hellinger, Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen, J. Reine Angew. Math. 136 (1909) 210–271.
[37] B.G. Lindsay, Efficiency versus robustness: the case for minimum hellinger distance and related methods, Ann. Statist. 22 (2) (1994) 1081–1114.
[38] P.R. Hansen, A. Lunde, Realized variance and market microstructure noise, J. Bus. Econom. Statist. 24 (2) (2006) 127–161.
[39] F. Liu, A.A. Pantelous, H.-J. von Mettenheim, Forecasting and trading high frequency volatility on large indices, Quant. Finance (2017) (in press).
[40] D.A. Hsieh, Chaos and nonlinear dynamics: application to financial markets, J. Financ. 46 (5) (1991) 1839–1877.
[41] R.K. Mishra, S. Sehgal, N.R. Bhanumurthy, A search for long-range dependence and chaotic structure in Indian Stock Market, Rev. Financ. Econ. 20 (2) (2011) 96–104.
[42] E.E. Peters, A chaotic attractor for the S&P 500, Financ. Anal. J. 47 (2) (1991) 55–62.
[43] W.-C. Chen, Nonlinear dynamics and chaos in a fractional-order financial system, Chaos Solitons Fractals 36 (5) (2008) 1305–1314.
[44] E. Maasoumi, J. Racine, Entropy and predictability of stock market returns, J. Econometrics 107 (1) (2002) 291–312.
[45] M.D. McKenzie, Chaotic behavior in national stock market indices: New evidence from the close returns test, Global Financ. J. 12 (1) (2001) 35–53.
[46] M.J. Hinich, D.M. Patterson, Evidence of nonlinearity in daily stock returns, J. Bus. Econom. Statist. 3 (1) (1985) 69–77.
[47] P. Anagnostidis, C.J. Emmanouilides, Nonlinearity in high-frequency stock returns: Evidence from the Athens Stock Exchange, Physica A 421 (2015) 473–487.
[48] A. Banerjee, J.J. Dolado, J.W. Galbraith, D. Hendry, Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data, OUP Catalogue, 1993.