ELSEVIER

Physica A 284 (2000) 348–354

PHYSICA A

www.elsevier.com/locate/physa

On the stock market recurrence

M.S. Baptista*, I.L. Caldas

Institute of Physics, University of São Paulo, C. P. 66318, CEP 05315-970 São Paulo, S.P., Brazil

Received 6 March 2000

Abstract

We analyze the return of the S & P 500 index and characterize its evolution as being typical of a low-dimensional recurrent deterministic system. The first Poincaré return time of the chaotic logistic mapping trajectories is used to model the return evolution. The efficiency of the model is demonstrated by daily predictions over an interval of time since January, 1950 of this index, and long-term prediction for a period of 150 days.

PACS: 00.05.45.Gg; 00.05.45.Tp; 80.89.90.+n

Keywords: Chaos; Econophysics; Stock Market; Dynamics; Modeling

1. Introduction

At present, there is no consensus about the dynamics of the stock market oscillations. Models used to derive stock quotes are based on empirical mathematics [1] which tries to mimic the single net return, defined as the rate that an asset changes its value over an interval of time. Variation of stock prices is usually considered a random process. From this point of view, there are different alternatives considering that the distribution of the return should be described by a normal distribution [2,3], a Lévy stable process [4–6], a leptokurtic distribution generated by a mixture of distributions [7], or by ARCH/GARCH models [8,9]. In Ref. [10], the authors speculate that the market dynamics could be turbulent, and thus higher dimensional, due to the fact that the return has a time scaling law similar to that found in the classical picture of turbulence by Kolmogorov. On the other hand, in Ref. [11], it is shown how to use multi-fractals in order to create a more realistic deterministic picture of the evolution of some market stocks.

* Corresponding author. Fax: +55-11-8186749.
E-mail address: murilo@if.usp.br (M.S. Baptista).

0378-4371/00/$ - see front matter
PII: S0378-4371(00)00226-0

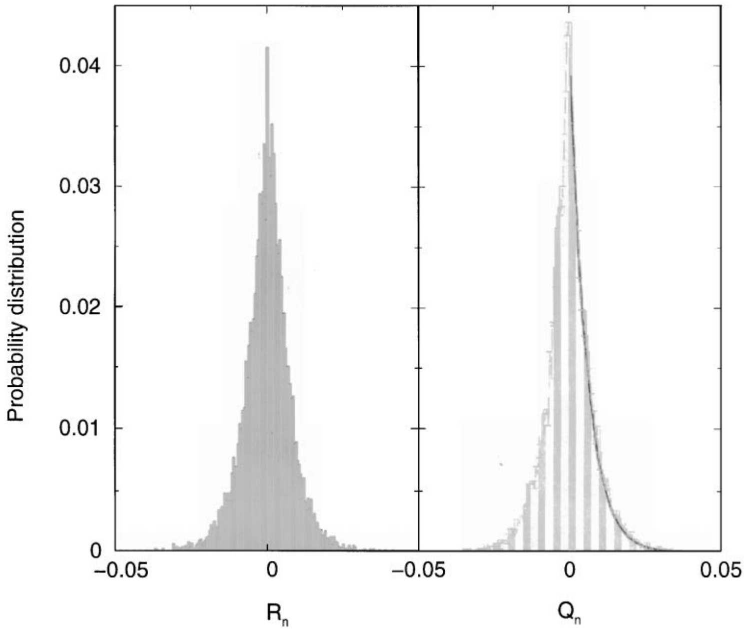

Fig. 1. (A) Distribution of the return R_n of the S & P index for a period from January 1950 up to October 1999. (B) Distribution of the simulated data generated by Eq. (5).

The purpose of this paper is to present a deterministic dynamical model for the return. We propose that the return is recurrent, i.e., it eventually comes back to the starting value (see Ref. [12]). Even though the returning time cannot be precisely determined, its average value over a large series of the returning time measurements, and other statistical variables like the variance of this distribution, are precisely determined. Consequently, our model is designed to mimic dynamically the evolution of any stock quote or index through the modeling of its return of a particular economic index. Here we use the Standard & Poor's 500 (S & P 500), Y_n, measured in time intervals, τ, spanning from 1 min up to two days. This article deals with the dynamics of stock market rather than specific risk analysis that will be considered in another publication.

The task of modeling the stock market boils down to the task of modeling the return of an asset, R_n, between the dates n and $n+1$, for a constant time interval τ, which is

$$R_n = \frac{Y_{n+1} - Y_n}{Y_n} . \tag{1}$$

Considering the daily evolution of S & P 500 since the year of 1950, deflated using the Consumer Price Index, the return gives the typical probability distribution $\rho(R(n))$ shown in Fig. 1 A. The function $\rho(R(n))$ has the same shape independently of τ. Two assertions about the distribution ρ have been proposed: It is non-Gaussian, and its variance should be finite. However, the determination of such distribution is currently a problem in econophysics. Recently, a proposal presented in Ref. [6] describes such distribution as being well modeled by a Lévy stable process. However, the Lévy distribution fails in describing the tails of the distribution.

In this work we present the function $\rho(R_n)$ as a sum of two Poisson distributions. Thus,

$$\rho(R_n) = \sigma_1 \exp\left(\frac{R_n - \langle R_n \rangle}{\langle R_n^+ \rangle}\right) + \sigma_2 \exp\left(\frac{-R_n - \langle R_n \rangle}{\langle R_n^+ \rangle}\right), \tag{2}$$

where for R_n positive, $\sigma_1 = 0$ and, for R_n negative, $\sigma_2 = 0$. Furthermore, for σ_1 and $\sigma_2 \neq 0$ they assume the same value proportional to $1/\langle R_n^+ \rangle$. Complementarily, we observe that

$$\langle R_n^+ \rangle \propto \tau^\alpha, \tag{3}$$

where R_n^+ is the R_n bigger than $\langle R_n \rangle$. Where $\langle R_n \rangle$ is the average of the R_n's.

Inspired by these facts, we propose a recurrent low-dimension dynamical model for the oscillation of the stock market assets. To model the stock market we look for a function Q_n that is equivalent to R_n, i.e., presenting the same statistics, given by a distribution g, and scaling properties verified in the stock market asset oscillations. So,

$$\rho(R_n) = g(Q_n) \tag{4}$$

and the model for the discrete evolution of the index for a time interval τ is

$$Z_{n+1} = Z_n(1 + Q_n)\,. \tag{5}$$

Note that Eq. (5) is Eq. (1) for $R_n = Q_n$ and $Y_n = Z_n$.

We conjecture that the return R_n can be modeled by a recurrent process. Thus, the proper type of system to model the return are the ergodic [13] ones, once all ergodic systems are recurrent. So, we model the return using the chaotic Logistic map,

$$x_{i+1} = 4x_i(1 - x_i)\,. \tag{6}$$

Further, it is shown how chaos theory is used to describe the nature of the stock market. From this theory, a few concepts are withdrawn from Eq. (6) as the existence and properties of the infinite number of periodic orbits embedded in its chaotic attractor, the existence of an invariant probability measure of the trajectories, and the correspondence of a few average quantities with its fractal dimension.

The relation between this chaotic dynamics and the return R_n is done through the Poincaré first return time $P_n(\varepsilon)$ of a chaotic trajectory, where $n = \{1, 2, \ldots, M\}$ is the number of times the trajectory of (6) falls in a x interval I of length ε, and P_n gives the time this trajectory takes to return to I. If x_0 belongs to I, and iterating (6) with x_0, the first point to fall in I is x_i, then the return time for $n = 1$ is $P_1 = i$. In this paper, we choose I as the interval $[0.100, 0.105]$. Numerically, we find that

$$\langle P_n(\varepsilon) \rangle \propto \varepsilon^{-0.99193}\,. \tag{7}$$

By Ref. [14] the average value $\langle P_n \rangle \propto \varepsilon^{-1/D_0}$, where D_0 is the fractal dimension (capacity dimension) of Eq. (6). The result of Eq. (7) agrees with this power law prediction because $D_0 \approx 1.017$ (computed with the box counting algorithm of Ref. [15]) for Eq. (6).

Eq. (7) lead us to scaling properties in R_n, representing the existence of a hidden dynamical process. If we assume that the interval length ε is related to τ such that $\varepsilon \propto 1/\tau$, we find that $\alpha = k/D_0$ in Eq. (3). Therefore, the hidden dynamics of the stock market should be low-dimensional once the dimension D_0 is close to one.

For a specific interval length ε, the probability distribution $g(P_n)$ is obtained by the knowledge that the number of periodic orbits, embedded in the chaotic attractor, that pass through the interval length ε, with period lower than T, is $N(T) \approx \exp hT$, where h is the Kolmogorov–Sinai entropy (see Ref. [16]). Thus, following Ref. [16], we obtain the Poisson distribution

$$g(P_n) = (1/\langle P_n \rangle)\exp(-P_n/\langle P_n \rangle)\,. \tag{8}$$

The distribution of a series composed by the linear combination of two Poincaré return time

$$Q_n = [aP_n - P_{n+1}]/F \tag{9}$$

gives for Q_n the same type of distribution of Eq. (2), where a is the shift of the distribution on $Q_n = 0$ and reflects the asset rise/fall tendency whether $a > 1$ or $a < 1$ over a given period of time $m\tau$, and F is a scaling factor. In order to have the value given by Q_n in the same units of R_n, and so $\rho(R_n) = g(Q_n)$, as required by Eq. (4), we use the scaling factor F, where F is a function of $m\tau$ and ε, calculated analyzing the m previous values $R_n(\tau)$ in the past, with $m \leqslant n$. Choosing a and F to satisfy $\langle Q_n^+ \rangle = \langle R_n^+ \rangle$ (Q_n^+ is the average of the Q_n's that are bigger than $\langle Q_n \rangle$) and $\langle Q_n \rangle = \langle R_n \rangle$, we obtain

$$a = \frac{1}{1 - \langle R_n \rangle / \langle R_n^+ \rangle} \tag{10}$$

and

$$F = a\frac{\langle P_n \rangle}{\langle R_n^+ \rangle}\,. \tag{11}$$

$\langle R_n^+ \rangle$ and $\langle R_n \rangle$ are the average values calculated for an assembling of the m successive previous values. $\langle P_n \rangle$ is given by Eq. (7).

To obtain Eqs. (10) and (11) we use the fact that the probability measure of data from Eq. (9) given by $\int g(Q_n)\,\mathrm{d}Q_n$ is μ/F, where μ defined by $\mu = \int g(P_n)\,\mathrm{d}P_n$ is invariant, it does not depend on the initial condition x_0 of Eq. (6). Thus, from Eq. (9) and as a consequence of the invariability of the probability measure of Q_n and P_n, $\langle P_n \rangle = \langle P_{n+1} \rangle$, $\langle Q_n \rangle = (a-1)\langle P_n \rangle / F$, and $Q_n^+ = \langle Q_n \rangle + \langle P_n \rangle / F$. Even though the series of Q_n's is obtained by a low-dimensional system, such numbers present stochastic-like properties and thus g is a distribution of a higher-dimensional set of numbers Q_n, and Eqs. (9)–(11) are derived from average statistical analysis.

Considering the period of time $\tau = 1$ day, $m = 12\,592$ values, relative to the S & P 500, $\langle R_n^+ \rangle = 0.00586911915$, $\langle R_n \rangle = 0.000239843459$. So, using Eqs. (10) and (11) for $\langle P_n \rangle = 189.8102$, $a = 1.04260645$ and $F = 33718.4058$. As expected from Eq. (9) $\langle Q_n \rangle = 0.000243619956$ and $\langle Q_n^+ \rangle = 0.00592876648$. Fig. 1B shows the distribution of the data generated by Eq. (5) iterated 12 592 times, for any typical initial condition. We

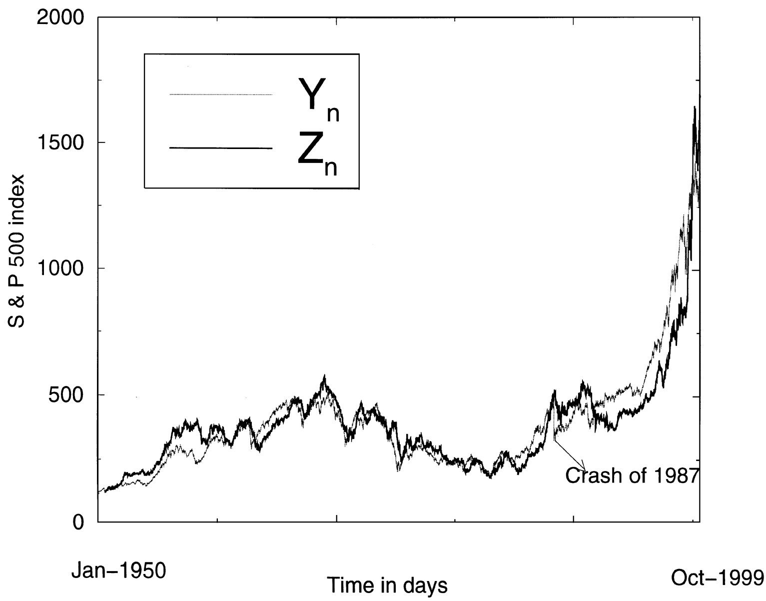

Fig. 2. The index S & P 500, since January 1950 till October 1999 is shown with a dashed line. The thin line shows a one day prediction for all the days of this period. Z_n (for a previous time $m\tau = 150$ d) is changed only in the beginning of the prediction.

see that the index return Y_n and the modeled index return Z_n have the same statistics. Also in this figure, we fit for values Q_n^+ a Poisson of form (8), and we find that the term inside the exponential is 179.15, which is a value very close to $1/\langle Q_n^+ \rangle$ as it is expected, once $g(Q_n^+)$ is a distribution of the form of Eq. (8).

In order to use the proposed model to make a prediction that corresponds to a possible evolution of a certain index, or any stock (we have shown that the model is suitable to a large number of different stocks), two parameters must be adjusted. Those are a and F. We keep ε constant, and set $m = 150$, and $\tau = 1$ day. Thus, to obtain a one day prediction using (5), we set $Z_{150} = Y_{150}$ and obtain Z_{151} by calculating F and a, using (11) and (10), respectively, where $\langle R_n^+ \rangle$ and $\langle R_n \rangle$ are calculated considering a set of m past data: R_{150}, $R_{150-1}, R_{150-2}, \ldots, R_1$. To obtain the prediction for the day 152, we calculate a and F from the set of data $R_{151}, R_{151-1}, \ldots, R_2$, and so on. In Fig. 2 we show the index S & P, since January 1950, with a dashed line. The thin full line shows a one day prediction for the whole data (from 1950 up to 1999). We clearly see how good the model is (5). We do not exclude for the calculations of a and F the data corresponding to the crash of 1987. Therefore, we see that the model adjusts itself for the index value after the crash. In fact, disregard of this period affects prediction for the period after the crash. The determination of a proper value for $m(t)$ for keeping the desired predictability, depends on the economic situation and has still to be explored.

For long-term prediction, we calculate the parameters a and F, using a high value for m. To predict the index in the day $p+1$, throughout a 150 days period (with p varying

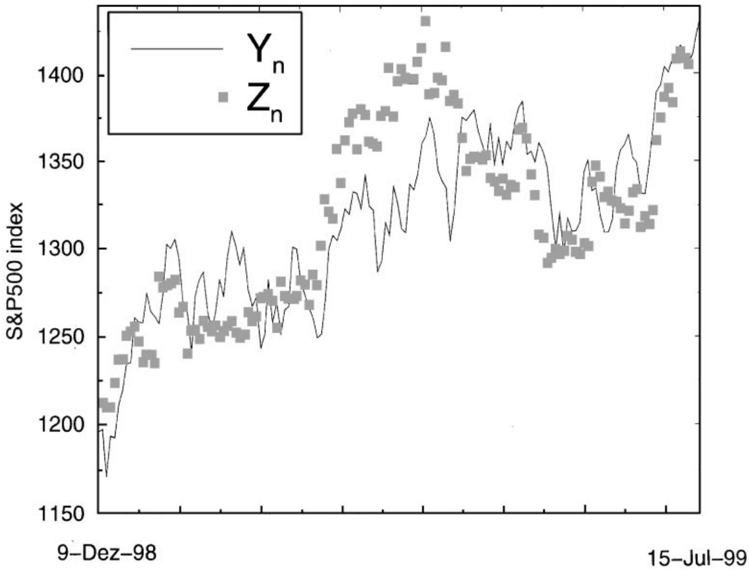

Fig. 3. One hundred and fifty days prediction of the index S & P 500 from December 9, 1998 until July 16, 1999, using Eq. (5) shown by the filled boxes and the S & P index shown with thin line (for a previous time $m\tau = 1000$ days).

from 1 to 149), a and F are calculated considering the set R_{p-m}, R_{p-m-1}, …, Q_{p-1}, Q_p. Thus, the previously predicted index values are used for future predictions. In Fig. 3, we predict the index from December 9, 1998 until July 16, 1999, using Eq. (5). We set $m = 1000$. We emphasize that Z_0 in Figs. 2 and 3 is changed only in the beginning of the prediction.

Another way of prediction is by averaging a few sets of solutions of Eq. (5), each one computed for either different values m, or for different values x_0. However, this is subject for further exploration.

We illustrated the model for $\tau = 1$ day. However, any other interval of time can be considered. In fact, a power-law scale of $\langle R_n^+ \rangle$ with τ can be obtained, and this would correspond to a signature of the particular index or stock. This scaling would give us enough information to adjust model (5) in order to start making predictions.

We have found that $\langle P_n^m \rangle$ scales with $\tau^{m/3}$, like observed in Ref. [10]. Due to the invariability of the distribution $g(P_n)$ for any initial condition, we emphasize that the non-linear function used to obtain the first Poincaré time return can be any, with the condition that is ergodic. ε can be adjusted to obtain a particular Q_n by Eq. (7), and the initial condition can be arbitrary. Q_n can be treated as being almost a series of random numbers, as currently believed, but in fact, these numbers were generated by a dynamical process.

We finish by concluding that the market is *dynamically recurrent* and a combination of the first Poincaré return time is a suitable model to predict future evolutions of some market index or stock.

Acknowledgements

The first author would like to thank Mauro C. Baptista, Marcos C. Baptista, and Marcelo C. Baptista for useful discussions and Kathryn A. Hebble for helping in revising the manuscript. Part of the data used for this work was downloaded from the Web site http://finance.yahoo.com/. This work is partially supported by FAPESP and CNPq.

References

[1] J.Y. Campbell, A.W. Lo, A.C. MacKinlay, The Econometrics of Financial Markets, Princeton University Press, Princeton, 1997.
[2] L.J.B. Bachelier, Theory the la Speculation, Gauthier-Villars, Paris, 1900.
[3] M.F.M. Osborne, Oper. Res. 7 (1959) 145.
[4] B.B. Mandelbrot, J. Bus. 36 (1963) 394.
[5] E.F. Fama, J. Bus. 38 (1965) 34.
[6] R.N. Mantegna, H.E. Stanley, Nature 376 (1995) 46.
[7] P.K. Clark, Econometrica 41 (1973) 135.
[8] R.F. Engle, Econometrica 50 (1982) 987.
[9] T. Bollerslev, R.Y. Chou, K.F. Kroner, J. Econom. 52 (1992) 5.
[10] S. Ghashghhaie, et al., Nature 381 (1996) 767.
[11] B.B. Mandelbrot, Sci. Am. (1999) 50.
[12] M. Kac, Bull. Am. Math. Soc. 53 (1947) 1002.
[13] J. Guckenheimer, P. Holmes, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Springer, New York, 1997.
[14] V. Afraimovich, Chaos 7 (1997) 12.
[15] H.E. Nusse, J.A. Yorke, Dynamics: Numerical Explorations, Springer, New York, 1997.
[16] G.M. Zaslavsky, M.K. Tippett, Phys. Rev. Lett. 67 (1991) 3251.