Eur. Phys. J. B **40**, 347–352 (2004)
DOI: 10.1140/epjb/e2004-00269-7

The European Physical Journal B

Conditional dynamics driving financial markets

M. Boguñá[a] and J. Masoliver

Departament de Física Fonamental, Universitat de Barcelona, Avinguda Diagonal 647, 08028 Barcelona, Spain

Received 28 October 2003 / Received in final form 14 April 2004
Published online 31 August 2004 –

Abstract. We revisit the problem of daily correlations in speculative prices and report empirical evidences on the existence of what we term a conditional or dual dynamics driving the evolution of financial assets. This dynamics is detected in several markets around the world and for different historical periods. In particular, we have analyzed the DJIA database from 1900 to 2002 as well as 65 companies trading in the LIFFE market of futures and 12 of the major European and American treasury bonds. In all cases, we find a twofold dynamics driving the financial evolution depending on whether the previous price went up or down. We conjecture that this effect is universal and intrinsic to all markets.

PACS. 89.65.Gh Economics; econophysics, financial markets, business and management – 05.45.Tp Time series analysis – 87.23.Ge Dynamics of social systems

1 Introduction

One fundamental assumption lying behind many modern theories of mathematical finance is the so called "efficient market hypothesis" which basically states that the market incorporates instantaneously any information concerning future market evolution [1]. In consequence, if a market is efficient with respect to some information set it is impossible to make economic profits by trading on the basis of that information set [2]. This in particular indicates that market efficiency necessarily implies the absence of (auto)correlations in financial prices at any time scale, for correlation means some degree of predictability which in turn would open the door to profitable strategies exclusively based on the information contained in the price itself. Note incidentally that the lack of correlations implied by the efficient market hypothesis means that the price process must be driven by white noise. However, this assumption is very restrictive since real markets are not efficient, at least at short times, and the existence of correlations seems to be well documented [2–6].

Therefore, the search for correlations in financial time series has been the subject of intense research during the last years [2–12]. Partly due to the hope that this knowledge would be useful for predicting the behavior of the market and, in a more academic sense, because correlations could bring some light to the understanding of the real market dynamics.

Among such correlations we want to single out the following one: it has been observed that the logarithmic variations of price –the so called returns– are correlated with themselves in such a way that highly positive returns are followed by highly positive returns as well. In the economics literature this effect applied to daily returns has been known for long. Thus Fama in 1970 observed slightly positive autocorrelations in daily security returns with a lag of one day and no evidence for higher lags [1]. The study of this particular and relevant correlation was pursued over the years by a number of authors who basically tried to find an explanation of it which essentially relies on non synchronous trading (see, for instance, Atchison et al. [11]). In the last decade the problem was again taken up by LeBaron who thoroughly studied daily and weakly serial correlations and confirmed Fama's findings of small but significant autocorrelation in returns for one day and no significance for higher lags [12] (see also [13] for a thorough review). This effect was applied by the author himself as a possible forecasting tool, unfortunately with tiny improvements [14].

In the econophysics literature the effect has also been considered and recently reported but only for high frequency data [15,16]. Nevertheless, this correlation is found to be of short range since it decays exponentially with a characteristic time of the order of minutes [8,9,15,16].

In this paper we want to revisit the daily correlation in returns and look at the problem from a different perspective, more from the viewpoint of econophysics and its methodology than that of mathematical finance. The paper is organized as follows. In Section 2 we analyze the signs of the returns. In Section 3 we present a simple two-state model which accounts for the main features of the effect. In Section 4 we address the problem of the universality of the conditional dynamics by showing that it

[a] e-mail: mbogunya@ffn.ub.es

Table 1. Summary of the empiric statistics for the DJIA index compared to the predictions of the uncorrelated model and the two-state model. The empty values in rows 3 and 4 must be taken from the empiric measurements given in row 2 as the inputs for the theoretical predictions of each model. Numbers within parentheses are affected by statistical errors (see main text).

	p_+	p_{++}	p_{--}	$\langle\tau_+\rangle$	$\langle\tau_-\rangle$	$\langle R_+\rangle$	$\langle R_-\rangle$	$\langle R\rangle$
Empiric DJIA index	0.52(2)	0.54(7)	0.49(5)	2.2(2)	2.0(2)	$8.(2)\times10^{-4}$	$-5.(3)\times10^{-4}$	$1.(7)\times10^{-4}$
Uncorrelated model	–	0.52(2)	0.47(8)	2.0(9)	1.9(1)	$1.(7)\times10^{-4}$	$1.(7)\times10^{-4}$	–
Two-state model	0.52(7)	–	–	2.2(1)	1.9(8)	–	–	$1.(8)\times10^{-4}$

appears in a wide range of corporate stocks and treasury bonds besides market indices. Conclusions are drawn in Section 5.

2 Signs of the return

Let us address the problem by first considering the time series of the sign of daily returns, R_n, defined as

$$R_n = \ln\left[S_n/S_{n-1}\right], \tag{1}$$

where S_n and S_{n-1} are the closing market prices corresponding to days n and $n-1$ respectively. We thus assign the value $+1$ if a given day has a positive return and -1 otherwise. Which are the statistical properties of this signal? The simplest hypothesis would be to consider positive and negative days as uncorrelated random events with a probability p_+ for positive days and $1-p_+$ for negative ones. Using this assumption as a test or "null hypothesis", the probability of having a sequence of n consecutive positive returns, $\psi_+(n)$, is therefore

$$\psi_+(n) = p_+^{n-1}(1-p_+), \tag{2}$$

that is, the geometric distribution. The average number of consecutive positive days is simply given by $\langle\tau_+\rangle = (1-p_+)^{-1}$ and the same holds for sequences of days with negative returns replacing p_+ by $1-p_+$, i.e., $\langle\tau_-\rangle = p_+^{-1}$.

In order to accept or reject the null hypothesis, we use data from the Dow Jones Industrial Average index (DJIA) containing daily records from 1900 to 2002 (28126 days) which covers a wide temporal range with many different economic and political situations thus ensuring statistical significance. Direct measurements on this database yield for the frequency of positive days the value $p_+ = 0.522 \pm 0.002$ and, according to the model given by equation (2), the expected number of consecutive positive and negative days is $\langle\tau_+\rangle_{theoretical} = 2.09 \pm 0.01$ and $\langle\tau_-\rangle_{theoretical} = 1.91 \pm 0.01$ respectively (these results are summarized in Tab. 1).

Figure 1 shows the probability distributions of the lengths of sequences of positive and negative days, $\psi_+(n)$ and $\psi_-(n)$. As is clearly seen, these distributions seem to follow a geometric law, in agreement with equation (2). However, the empirical average lengths of positive and negative days obtained from direct measurements are $\langle\tau_+\rangle_{empiric} = 2.22 \pm 0.02$ and $\langle\tau_-\rangle_{empiric} = 2.02 \pm 0.02$ which are higher than the theoretical values predicted above. The disagreement between empiric and theoretical results is certainly small and might go easily unnoticed, although a careful analysis of these values which takes into account the statistical errors leads to the rejection of the null hypothesis (2) of independence of positive and negative returns.

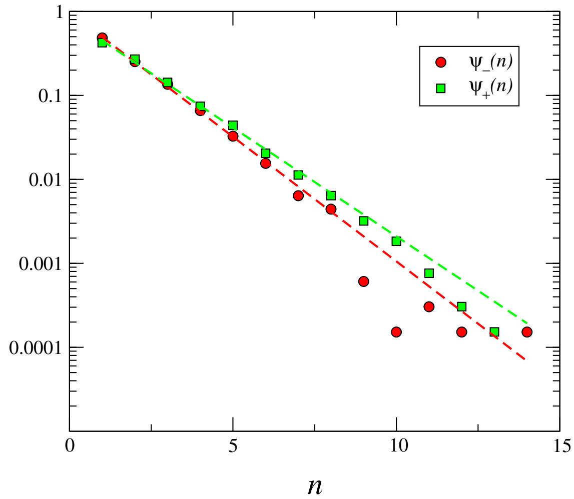

Fig. 1. Probability of having a sequence of n consecutive positive or negative days. The solid lines are the geometric distributions discussed in the text with average values given by $\langle\tau_+\rangle_{empiric} = 2.22$ and $\langle\tau_-\rangle_{empiric} = 2.02$.

These results suggest that markets behave differently whenever there is a sequence of positive or negative returns. On the other hand, the geometric form for the distribution of lengths of those sequences (which in the continuous limit yields the Poisson distribution) indicates that the market is Markovian. Therefore, no information can be extracted from the elapsed time since the last change of sequence and, therefore, the memory of the market must be, at most, of one single day. This implies that the return of the price during a given day can only be correlated with the previous day, in particular with the sign of the previous day.

3 The two-state model

Perhaps one of the simplest model still able to reproduce all of the above empirical observations is a two-state model in which the probability of having positive or negative returns depends on the sign of the previous day.

More precisely, let p_{++} be the probability of having a positive return given that the return of the previous day was positive and p_{--} the probability of having a negative return given that the return of the previous day was negative. Note that the model has only two independent parameters, p_{++} and p_{--}. The rest of probabilities can be obtained from them as $p_{-+} = 1-p_{++}$ and $p_{+-} = 1-p_{--}$ and they measure the probability of having a negative (positive) return given that the return of the previous day was positive (negative).

The distributions $\psi_+(n)$ and $\psi_-(n)$ are now given by

$$\psi_+(n) = p_{++}^{n-1}(1-p_{++}) \tag{3}$$

and

$$\psi_-(n) = p_{--}^{n-1}(1-p_{--}). \tag{4}$$

Again geometric distributions with average lengths given by

$$\langle\tau_+\rangle = (1-p_{++})^{-1} \quad \text{and} \quad \langle\tau_-\rangle = (1-p_{--})^{-1}. \tag{5}$$

The frequency of positive days, p_+, is easily evaluated by observing that in a two-state system with states + and − the frequency of each state is

$$p_{+,-} = \frac{\langle\tau_{+,-}\rangle}{\langle\tau_+\rangle + \langle\tau_-\rangle} \tag{6}$$

independent of the waiting time distributions [17]. Substituting equation (5) into equation (6) yields

$$p_+ = \frac{1-p_{--}}{2-p_{--}-p_{++}}, \tag{7}$$

and a similar expression for p_- exchanging p_{++} by p_{--}.

We will now compare the predictions of the two-state model with empirical data. Direct measurements on the DJIA data set yield the following values for the conditional probabilities: $p_{++} = 0.547 \pm 0.004$ and $p_{--} = 0.495 \pm 0.004$. Using these two measures as inputs for the model, the predicted values for p_+, $\langle\tau_+\rangle$, and $\langle\tau_-\rangle$ are $p_+ = 0.527 \pm 0.004$, $\langle\tau_+\rangle = 2.21 \pm 0.02$ and $\langle\tau_-\rangle = 1.98 \pm 0.02$, in perfect agreement with the empirical results reported above (see Tab. 1).

It might be argued that the discrepancy between the empirical measures of $\langle\tau_+\rangle$ and $\langle\tau_-\rangle$ and the theoretical predictions of the uncorrelated model are marginal and, consequently, the two-state model only introduces a slight correction to the actual dynamics. Nevertheless, what seems to be significant is the fact the market apparently reacts differently depending on the sign of the previous day and this introduces, in a natural way, the idea of a dual dynamics. Having this in mind, we define $p(R|R_{prev} > 0)dR$ to be the conditional probability that the daily return lies within the interval $(R, R+dR)$ given that the previous day had a positive return. Analogously $p(R|R_{prev} < 0)dR$ is that conditional probability if the previous day had a negative return. Up to this point, we have only studied the behavior of the sign of the signal specified by the quantities p_{++} and p_{--}, which are related to the previous functions by

$$p_{++} = \int_0^\infty p(R|R_{prev} > 0)dR \tag{8}$$

$$p_{--} = \int_{-\infty}^0 p(R|R_{prev} < 0)dR. \tag{9}$$

However, if the market is really driven by a dual dynamics there should be a substantial difference between the moments of $p(R|R_{prev} > 0)$ and $p(R|R_{prev} < 0)$. Let us denote by $\langle R_+\rangle$ and $\langle R_-\rangle$ the first moment of these distributions, that is, $\langle R_+\rangle$ [$\langle R_-\rangle$] is the conditional average of the daily return given that yesterday's return was positive [negative]. Similarly, let σ_+, and σ_- be their standard deviations. For the DJIA index, the empirical values for these quantities are: $\langle R_+\rangle = (8.2 \pm 0.8) \times 10^{-4}$, $\langle R_-\rangle = (-5.3 \pm 1.0) \times 10^{-4}$, $\sigma_+ = (9.9 \pm 0.2) \times 10^{-3}$, and $\sigma_- = (11.8 \pm 0.3) \times 10^{-3}$. These values should be compared with the unconditional average of the daily return, $\langle R\rangle = (1.7 \pm 0.6) \times 10^{-4}$, and volatility $\sigma = (10.9 \pm 0.2) \times 10^{-3}$. Note that $\langle R\rangle$ and its variance can be evaluated through the two-state model by

$$\langle R\rangle = p_+\langle R_+\rangle + p_-\langle R_-\rangle, \tag{10}$$

and

$$\sigma = \sqrt{p_+\sigma_+^2 + p_-\sigma_-^2 + p_+p_-(\langle R_+\rangle - \langle R_-\rangle)^2}, \tag{11}$$

with the results $\langle R\rangle = (1.8 \pm 0.6) \times 10^{-4}$ and $\sigma = (10.9 \pm 0.2) \times 10^{-3}$. Both in very good agreement with their empirical values. Table 1 summarizes the relevant statistics for the DJIA index and the equivalent values predicted by the uncorrelated model and the two-state model.

There is something quite significant in these results, for they show that the average return of the market is the result of the composition of two independent signals: one of them positive, $\langle R_+\rangle$, and another one negative, $\langle R_-\rangle$. At the light of these results, and given the multiplicative character of the market, it does not seem to be possible to neglect the effects of this dual dynamics, at least in the long run. Indeed, the quantitative difference between the average daily return of both signals is rather significant in the sense that a small change in the signal would substantially alter the long term trend of the market (see Fig. 3).

As we have seen, the daily return of a given day is a random quantity correlated with the return of the previous day. One question that arises now is: how does this correlation depend on the magnitude of the previous return? In order to check this point we evaluate the average return given that the previous day had a return greater than a certain value r_c, $\langle R_+(r_c)\rangle$, or smaller than r_c, $\langle R_-(r_c)\rangle$. Note that $r_c = 0$ correspond to the previous analysis. These two functions are plotted in Figure 2. As is clearly seen, there is a significant difference whenever the previous day has a positive or negative increment in

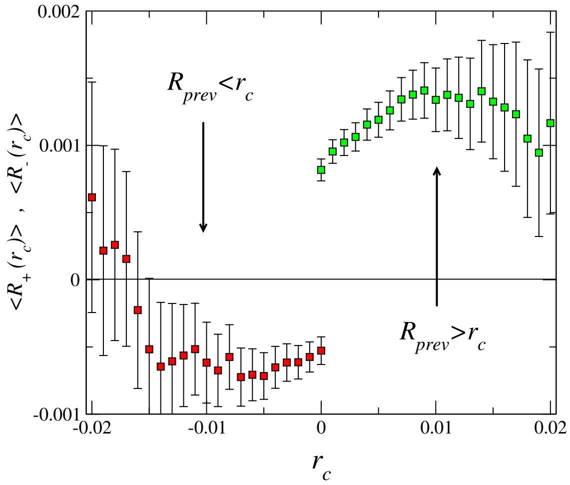

Fig. 2. Average daily return given that the previous day had a return greater than r_c (right) and given that the previous day has a return smaller than r_c (left).

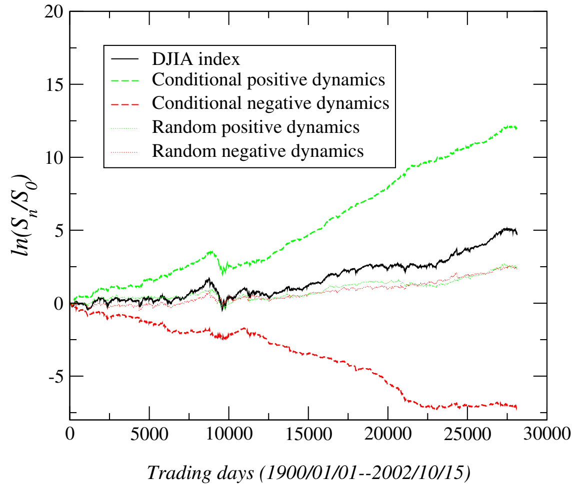

Fig. 3. Evolution of the logarithm of the DJIA index compared to the corresponding positive and negative dynamics along with a simulation of the expected dynamics if returns were uncorrelated. Observe that in both cases the sum of the two signals is equal to the logarithm of the index price.

price. For $r_c \in [-1.5\%, 1.5\%]$ the positive branch is positive –and slightly increasing – whereas the negative branch remains negative and almost insensitive to the magnitude of the previous price drop. Beyond this interval, the negative branch increases and, eventually, both branches become equivalent —within the statistical error— meaning that correlations are lost for this range of returns. In other words, there is no net effect if the previous day has a return greater than 1.5% or smaller than −1.5%. This reversion of the negative branch could be understood as a recovery effect after an extreme price drop, although the poor statistics for this range of returns does not allow us to make a more assertive and documented statement.

It is worth noticing that the effect described in Figure 2 is similar to that observed by LeBaron [12] who showed that serial correlations (measured through the correlation coefficient between two consecutive returns) are significantly bigger during low volatility periods than during highly volatile periods (see also [13]). Note, however, that the present approach essentially differs to that of [12] in the fact that we measure serial correlations by means of two independent signals (positive and negative returns) instead of the entire correlation coefficient which embodies the two signals. Moreover in [12] the correlation coefficient is conditioned to the volatility of the previous period evaluated on an arbitrary time window, which can introduce additional uncertainties.

As we have mentioned, the conditional dynamics is basically observed when previous returns lay in the interval $[-1.5\%, 1.5\%]$. For higher values of previous returns this twofold dynamics turns into the standard dynamics with no bias between positive and negative previous returns. For the DJIA index these highly volatile days account for less than 10% out of the total trading days. However, this 10% of days does not lessen the relevance of the correlations present in the remaining 90% of trading days. To enhance the relevance of this effect we will analyze the evolution of the index in terms of the dual dynamics. Thus the evolution of the price index at time n, S_n, can be expressed as

$$\ln\left[\frac{S_n}{S_0}\right] = \sum_{i=1}^{n} R_i = \sum_{R_{prev}>0} R_i + \sum_{R_{prev}<0} R_i, \qquad (12)$$

where the last two terms correspond to the sum of returns for which the previous return was positive or negative respectively. In Figure 3 we show the evolution of the logarithm of the DJIA index as well as the evolution of both branches of the conditional dynamics. As is clearly seen, the positive branch increases faster than the price itself whereas the negative one decreases. Figure 3 also shows the evolution of two signals constructed from a random partition of returns into two groups under the constrain that one of the groups has np_+ returns and the other $n(1-p_+)$. Note that these signals would be the expected ones if returns were uncorrelated, in other words, they represent a simulation of the unconditional dynamics. Observe the great divergence between the conditional dynamics and the unconditional one. We conjecture that this difference is mainly due to those returns with low previous returns since correlations are destroyed otherwise. Figure 3 also seems to discover another interesting property, namely, the independence of both branches of the dual dynamics. This is best seen in the evolution of the last 30 years (see Fig. 3 from day 22500 on) during this period the negative branch remains flat which seems to indicate the absence of any drift during days with previous negative returns. This, in turn, suggests that the net drift of the index would be predominately due to the positive branch.

4 Analysis for individual stocks and treasury bonds

We finally address the question of the universality of the dual dynamics. The preceding analysis has been carried out for one specific index, the DJIA, during a period of 100 years. Previous works on the effect have also been performed on daily data of DJIA and Standard and Poors indices [12]. Therefore, one important point is whether this correlation is also present for individual stocks and any other class of financial assets. In order to shed some light on this question, we have analyzed the performance of 65 companies trading in the LIFFE market[1] during a twelve year period from 1990 to 2002. In this case the increase of statistical error due to the short period considered is balanced by analyzing a large number of different companies. For each company we have measured $\langle R_\pm \rangle$, that is, the average conditional returns given that yesterday return was positive, $\langle R_+ \rangle$, or negative, $\langle R_- \rangle$. The results are shown in Figure 4 as a scattered plot,where axes represent the conditional average daily returns rescaled by the unconditional volatility of the corresponding company. If no correlation were present between a given return and the sign of the previous one then $\langle R_+ \rangle$ and $\langle R_- \rangle$ would take the same value (except for statistical fluctuations) and, therefore, all companies would be scattered around the main diagonal, in the first quadrant. In contrast we see in Figure 4 a clear tendency to stay in the second quadrant, including the statistical error, with $\langle R_+ \rangle$ being a positive quantity and $\langle R_- \rangle$ being a negative one (or close to zero). This means that, on average, the returns after a positive day outperform those that follow a negative day in agreement with the model presented. The same effect is observed in other classes of financial assets, such as treasury bonds[2] —blue symbols in Figure 4— or commodities (not reported here). All these results suggest the universality of the dual dynamics driving the evolution of financial markets.

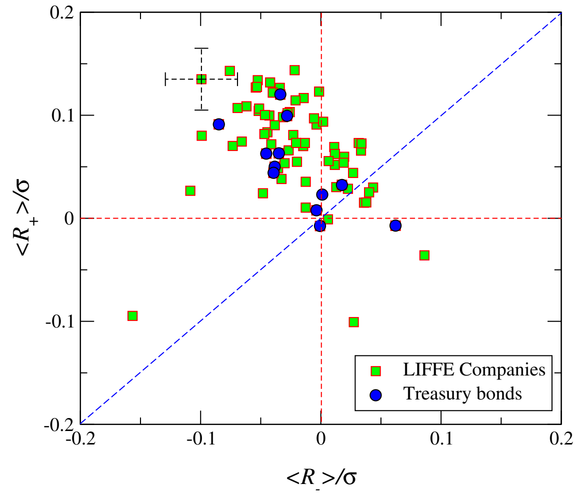

Fig. 4. Scattered plot of the average daily returns given that the previous day had a positive or negative increment for the companies trading in the LIFFE market and several American and European treasury bonds during the period 1990–2002. For the sake of comparison, these average returns are rescaled by the volatility of the corresponding company. The bars represent the average statistical error of all companies and bonds.

5 Conclusion

We have revisited the effect of correlations of daily returns and reported empirical evidences of the existence of a conditional dynamics driving the behavior of financial markets which manifests itself in the fact that daily prices tend to go up or down depending on whether yesterday's price went up or down. Moreover this dynamics seems to be ubiquitous to a wide sample of different markets which may indicate the universal character of this effect.

Let us now summarize the main correlations observed in financial time series which show the incompleteness of the efficient market hypothesis. For one hand, we have the return-return correlations which are observed in two different time scales: (i) the daily scale studied above and (ii) the high frequency scale with a correlation time of the order of few minutes [8,9,15,16]. A second important example is provided by the volatility-volatility correlation. In this case there seems to be a clear positive correlation between the dispersion of the return today and in the future, with a characteristic time of the order of years [5,8,9]. Finally, a third type of correlation is provided by the leverage effect [18,19], which states that a large drop of the price is followed by an increase of the volatility. This correlation is found to be of intermediate range, with a typical time scale of few weeks.

We close this paper by stressing the fact that financial time series are often non-stationary, at least at long times, and, consequently, it is possible to find short periods in which the dual dynamics is not clearly visible. Therefore, the empirical findings reported here must be considered from an overall point of view at the same level as the observation that the market is historically growing despite the existence of many bear periods. This point will be addressed in future communications.

We greatly acknowledge J.M. Porrà, J.-P. Bouchaud, and R. Pastor-Satorras for discussions and useful suggestions. We also thank two anonymous referees who indicated some important references. This work has been partially supported by DGI under Grant No. BFM2003-04574 and by GC under contract No. 2000 SGR-00023. M. B.acknowledges support from the European Commission FET Open Project No. COSIN IST-2001-33555 and the Ramon y Cajal Program of the Ministerio de Ciencia y Tecnologia (Spain).

[1] London International Financial Futures and Options Exchange.

[2] US Long Bond, US 10YR Note, US 5YR Note, US 2YR Note, US 3MO Treasury Bill, Euro-Bund Future, Euro-Bobl Future, Euro-Schatz Future, EUX 3 MO Euribor, Swiss Fed Bund Future, Euro Sfr 3 MO LIFFE, Long Gilt Future.

References

1. E. Fama, J. Finance **25**, 383 (1970)
2. J.Y. Campbell, A. Lo, A.C. MacKinlay, *The Econometrics of Financial Markets* (Princeton University Press, Princeton, 1997)
3. E. Fama, J. Finance **46**, 1575 (1991)
4. R. Mantegna, H.E. Stanley, *An introduction to Econophysics* (Cambridge University Press, Cambridge, UK 1999)
5. J.-P. Bouchaud, M. Potters, *Theory of Financial Risks* (Cambridge University Press, Cambridge, UK 2000)
6. Y.-C. Zhang, Physica A **269**, 30 (1999)
7. Z. Ding, C.W.J. Granger, R.F. Engle, J. Empirical Finance **1**, 83 (1993)
8. Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.-K. Peng, H.E. Stanley, Phys. Rev. E **60**, 1390 (1999)
9. P. Gopikrishnan, V. Plerou, L.A.N. Amaral, M. Meyer, H.E. Stanley, Phys. Rev. E **60**, 5305 (1999)
10. M. Potters, R. Cont, J.-P. Bouchaud, Europhys. Lett. **41**, 239 (1998)
11. M.D. Atchison, K.C. Butler, R.R. Simonds, J. Finance **42**, 111 (1987)
12. B. LeBaron, J. Business **65**, 199 (1992)
13. M.M. Dacorogna, R. Gençay, U.A. Müller, R.B. Olsen, O.V. Pictet, *An Introduction to High-Frequency Data* (Academic Press, San Diego, 2001)
14. B. LeBaron, J. Applied Econometrics **7**, 137 (1992)
15. T. Ohira, N. Sazuka, K. Marumo, T. Shimizu, M. Takayasu, H. Takayasu, Physica A **308**, 368 (2002)
16. N. Sazuka, T. Ohira, K. Marumo, T. Shimizu, M. Takayasu, H. Takayasu, Physica A **324**, 366 (2003)
17. M. Boguñá, A.M. Berezhkovskii, G.H. Weiss, Physica A **282**, 475 (2000)
18. J.-P. Bouchaud, A. Matacz, M. Potters, Phys. Rev. Lett. **87**, 228701 (2001)
19. J. Perelló, J. Masoliver, Phys. Rev. E **67**, 037102 (2003)