A dynamical model describing stock market price distributions

Jaume Masoliver[a,*], Miquel Montero[a], Josep M. Porrà[a,b]

[a]*Departament de Física Fonamental, Universitat de Barcelona, Diagonal, 647, Barcelona 08028, Spain*
[b]*Gaesco Bolsa, SVB, S.A., Diagonal, 429, Barcelona 08036, Spain*

Abstract

High-frequency data in finance have led to a deeper understanding on probability distributions of market prices. Several facts seem to be well established by empirical evidence. Specifically, probability distributions have the following properties: (i) They are not Gaussian and their center is well adjusted by Lévy distributions. (ii) They are long-tailed but have finite moments of any order. (iii) They are self-similar on many time scales. Finally, (iv) at small time scales, price volatility follows a non-diffusive behavior. We extend Merton's ideas on speculative price formation and present a dynamical model resulting in a characteristic function that explains in a natural way all of the above features. The knowledge of such a distribution opens a new and useful way of quantifying financial risk. The results of the model agree – with high degree of accuracy – with empirical data taken from historical records of the Standard & Poor's 500 cash index. © 2000 Published by Elsevier Science B.V. All rights reserved.

1. Introduction

One of the most important problems in mathematical finance is to know the probability distribution of speculative prices. In spite of its importance for both theoretical and practical applications the problem is yet unsolved. The first approach to the problem was given by Bachelier in 1900 when he modelled price dynamics as an ordinary random walk where prices can go up and down due to a variety of many independent

---

* Corresponding author.

E-mail address: jaume@ffn.ub.es (J. Masoliver)

random causes. Consequently, the distribution of prices was Gaussian [1]. The normal distribution is ubiquitous in all branches of natural and social sciences and this is basically due to the Central Limit Theorem: the sum of independent, or weakly dependent, random disturbances, all of them with finite variance, results in a Gaussian random variable. Gaussian models are thus widely used in finance although, as Kendall first noticed [2], the normal distribution does not fit financial data especially at the wings of the distribution. Thus, for instance, the probability of events corresponding to 5 or more standard deviations is around 10^4 times larger than the one predicted by the Gaussian distribution, in other words, the empirical distributions of prices are highly leptokurtic. Is the existence of too many of such events, the so-called outliers, the reason for the existence of "fat tails" and the uselessness of the normal density especially at the wings of the distribution. Needless to say, the tails of the price distributions are crucial in the analysis of financial risk. Therefore, obtaining a reliable distribution has deep consequences from a practical point of view [3,4].

One of the first attempts to explain the appearance of long tails in financial data was made by Mandelbort in 1963 [5] who, based on Pareto–Lévy stable laws [6], obtained a leptokurtic distribution. Nevertheless, the price to pay is high: the resulting probability density function has no finite moments, except the first one. This is indeed a severe limitation and it is not surprising since Mandelbrot's approach can still be considered within the framework of the Central Limit Theorem, that is, the sum of independent random disturbances of infinite variance results in the Lévy distribution which has infinite variance [6]. On the other hand, the Lévy distribution has been tested against data in a great variety of situations, always with the same result: the tails of the distribution are far too long compared with actual data. In any case, as Mantegna and Stanley have recently shown [7], the Lévy distribution fits very well the center of empirical distributions – much better than the Gaussian density – and it also shares the scaling behavior shown in the data [7–10].

Therefore, if we want to explain speculative price dynamics as a sum of weakly interdependent random disturbances, we are confronted with two different and in some way opposed situations. If we assume finite variance, the tails are "too thin" and the resulting Gaussian distribution only accounts for a narrow neighborhood at the center of the distribution. On the other hand, the assumption of infinite variance leads to the Lévy distribution which explains quite well a wider neighborhood at the center of distribution but results in "too fat tails". The necessity of having an intermediate model is thus clear and this is the main objective of this paper.

Obviously, since the works of Mandelbrot [5] and Fama [11] on Lévy distributions, there have been several approaches to the problem, some of them applying cut-off procedures of the Lévy distribution [12,13] and, more recently, the use of ARCH and GARCH models to obtain leptokurtic distributions [14]. The approaches based on cut-off procedures are approximations to the distributions trying to better fit the existing data, but they are not based on a dynamical model that can predict their precise features. On the other hand, ARCH [15] and GARCH [16] models are indeed dynamical adaptive models but they do not provide an overall picture of the market

dynamics resulting in a distinctive probability distribution. In fact, ARCH/GARCH models usually assume that the market is Gaussian with an unknown time-varying variance so as to be self-adjusted to obtain predictions.

The paper is organized as follows. In Section 2 we proposed that stochastic model and set the mathematical framework that leads to a probability distribution of prices. In Section 3 we present the main results achieved by the model. Conclusions are drawn in Section 4.

2. Analysis

Let $S(t)$ be a random process representing stock prices or some market index value. The usual hypothesis is to assume that $S(t)$ obeys a stochastic differential equation of the form

$$\frac{\dot{S}}{S} = \rho + F(t) , \tag{1}$$

where ρ is the instantaneous expected rate of return and $F(t)$ is a random process with specified statistics, usually $F(t)$ is zero-mean Gaussian white noise, $F(t) = \zeta(t)$, in other words $dW(t) = \zeta(t)\,dt$, where $W(t)$ is the Wiener process or Brownian motion. In this case, the dynamics of the market is clear since the return $R(t) \equiv \log[S(t)/S(0)]$ obeys the equation $\dot{R} = \rho + \zeta(t)$ which means that returns evolve like an overdamped Brownian particle driven by the "inflation rate" ρ and, in consequence, the return distribution is Gaussian.

Let us take a closer look at the price formation and dynamics. Following Merton [17], we say that the change in the stock price (or index) is basically due to the random arrival of new information. This mechanism is assumed to produce a marginal change in the price and it is modelled by the standard geometric Brownian motion defined above. In addition to this "normal vibration" in price, there is an "abnormal vibration" basically due to the (random) arrival of important new information that has more than a marginal effect on price. Merton models this mechanism as a jump process with two sources of randomness: the arrival times when jumps occur, and the jump amplitudes. The result of the overall picture is that the noise source $F(t)$ in price equation is now formed by the sum of the two independent random components

$$F(t) = \xi(t) + f(t) , \tag{2}$$

where $\xi(t)$ is Gaussian white noise corresponding to the normal vibration, and $f(t)$ is "shot noise" corresponding to the abnormal vibration in price. This shot noise component can be explicitly written as

$$f(t) = \sum_{k=1}^{\infty} A_k \delta(t - t_k) , \tag{3}$$

where $\delta(t)$ is the Dirac delta function, A_k are jump amplitudes, and t_k are jump arrival times. It is also assumed that A_k and t_k are independent random variables with known probability distributions given by $h(x)$ and $\psi(t)$, respectively [18].

We now go beyond this description and specify the "inner components" of the normal vibration in price, by unifying this with Merton's abnormal component. We thus assume that *all changes in the stock price* (*or index*) *are modelled by different shot-noise sources corresponding to the detailed arrival of information*, that is, we replace the total noise $F(t)$ by the sum

$$F(t) = \sum_{n=n_0}^{m} f_n(t) \ , \tag{4}$$

where $f_n(t)$ is a set of independent shot-noise processes given by

$$f_n(t) = \sum_{k_n=1}^{\infty} A_{k_n,n} \delta(t - t_{k_n,n}) \ . \tag{5}$$

The amplitudes $A_{k_n,n}$ are independent random variables with zero mean and probability density function (pdf), $h_n(x)$, depending only on a single "dimensional" parameter which, without loss of generality, we assume to be the standard deviation of jumps σ_n, i.e.,

$$h_n(x) = \sigma_n^{-1} h(x\sigma_n^{-1}) \ . \tag{6}$$

We also assume that the occurrence of jumps is a Poisson process, in this case shot noises are Markovian, and the pdf for the time interval between jumps is exponential:

$$\psi(t_{k_n,n} - t_{k_n-1,n}) = \lambda_n \exp[-\lambda_n(t_{k_n,n} - t_{k_n-1,n})] \ , \tag{7}$$

where λ_n are mean jump frequencies, i.e., $1/\lambda_n$ is the mean time between two consecutive jumps [18]. Finally, we order the mean frequencies in a decreasing way: $\lambda_n < \lambda_{n-1}$.

Let $X(t)$ be the zero-mean return, i.e., $X(t) \equiv R(t) - \rho t$. For our model $X(t)$ reads as

$$X(t) = \sum_{n=n_0}^{m} \sum_{k_n=1}^{\infty} A_{k_n,n} \theta(t - t_{k_n,n}) \ , \tag{8}$$

where $\theta(t)$ is the Heaviside step function. Our main objective is to obtain an expression for the pdf of $X(t)$, $p(x,t)$, or equivalently, the characteristic function (cf) of $X(t)$, $\tilde{p}(\omega,t)$, which is the Fourier transform of the pdf $p(x,t)$. Note that $X(t)$ is a sum of independent jump processes, this allows us to generalize Rice's methods for a single Markov shot noise to the present case of many shot noises [19]. The final result is

$$\tilde{p}(\omega,t) = \exp\left\{ -t \sum_{n=n_0}^{m} \lambda_n[1 - \tilde{h}(\omega\sigma_n)] \right\} \ . \tag{9}$$

As it is, $X(t)$ represents a shot-noise process with mean frequency of jumps given by $\lambda = \sum \lambda_n$ and jump distribution given by $h(x) = \sum \lambda_n h_n(x)/\lambda$. Nevertheless, we make a further approximation by assuming that (i) $n_0 = -\infty$, i.e., there is an infinite number of shot-noise sources, and (ii) there is no characteristic time scale limiting the maximum

feasible value of jump frequencies, thus, $\lambda_n \to \infty$ as $n \to -\infty$. Both assumptions are based on the fact that the "normal vibration" in price is formed by the addition of (approximately) infinitely many random causes, which we have modelled as shot noises. According to this, we introduce a "coarse-grained" description and replace the sum in Eq. (9) by an integral as

$$\tilde{p}(\omega, t) = \exp\left\{-t \int_{-\infty}^{u_m} \lambda(u)[1 - \tilde{h}(\omega\sigma(u))]\, du\right\}. \tag{10}$$

In order to proceed further we should specify a functional form for $\lambda(u)$ and $\sigma(u)$. We note by empirical evidence that the bigger a sudden market change is, the longer is the time we have to wait until we observe it. Therefore, since $\lambda(u)$ decreases with u (recall that frequencies are decreasingly ordered) then $\sigma(u)$ must increase with u. We thus see that $\sigma(u)$ has to be a positive definite, regular and monotonically increasing function for all u. The simplest choice is $\sigma(u) = \sigma_0 e^u$. On the other hand, there is empirical evidence of scaling properties in financial data [7–10]. We summarize the above requirements (i.e., inverse relation between λ and σ, and scaling) by imposing the "dispersion relation":

$$\lambda = \lambda_0(\sigma_0/\sigma)^\alpha. \tag{11}$$

where α is the scaling parameter. Under these assumptions, the cf of the return $X(t)$ reads:

$$\tilde{p}(\omega, t) = \exp\left\{-\lambda_0 t \sigma_0^\alpha \int_0^{\sigma_m} z^{-1-\alpha}[1 - \tilde{h}(\omega z)]\, dz\right\}, \tag{12}$$

where $\sigma_m = \sigma_0 e^{u_m}$ is the maximum value of the standard deviation. We observe that if $\sigma_m = \infty$, which means that some shot-noise source has infinite variance, then Eq. (12) yields the Lévy distribution

$$\tilde{L}_\alpha(\omega, t) = \exp(-kt\omega^\alpha), \tag{13}$$

where

$$k = \lambda_0 \sigma_0^\alpha \int_0^\infty z^{-1-\alpha}[1 - \tilde{h}(z)]\, dz. \tag{14}$$

Hence, if we want a distribution with finite moments, we have to assume a finite value for σ_m.

Let λ_m be the mean frequency corresponding to the maximum (finite) variance. Recall that, in the discrete case (c.f. Eq. (9)), shot-noise sources are ordered, thus, λ_m and σ_m correspond to the mean frequency and the variance of the last jump source considered. Our last assumption is that the total number of noise sources in Eq. (8) increases with the observation time t and, since $n_0 = -\infty$, this implies that $m = m(t)$ is an increasing function of time. Consequently, the mean period of the last jump source, λ_m^{-1}, also grows with t. The simplest choice is the linear relation: $\lambda_m t = a$, where $a > 0$ is constant. Therefore, from the dispersion relation, Eq. (11), we see that the maximum jump variance depends on time as a power law

$$\sigma_m^2 = (bt)^{2/\alpha}, \tag{15}$$

where $b \equiv \sigma_0^\alpha \lambda_0 / a$. We finally have

$$\tilde{p}(\omega, t) = \exp\left\{ -abt \int_0^{(bt)^{1/\alpha}} z^{-1-\alpha}[1 - \tilde{h}(\omega z)]\,\mathrm{d}z \right\}. \tag{16}$$

3. Results

Let us now present the main results and consequences of the above analysis. First, the volatility of the return is given by

$$\langle X^2(t) \rangle = \frac{a\sigma_m^2}{2 - \alpha} = \frac{a}{2 - \alpha}(bt)^{2/\alpha}, \tag{17}$$

which proves that $\alpha < 2$ and the volatility shows super-diffusion. The anomalous diffusion behavior of the empirical data (at least at small time scales) was first shown by Mantegna and Stanley [20,21]. Second, kurtosis is constant and is given by

$$\gamma_2 = \frac{(2 - \alpha)^2 \tilde{h}^{(iv)}(0)}{(4 - \alpha)a}. \tag{18}$$

Thus, $\gamma_2 > 0$ for all t, in other words, we have a leptokurtic distribution in all time scales. Third, the return probability distribution scales as

$$p(x, t) = (bt)^{-1/\alpha} p(x/(bt)^{1/\alpha}) \tag{19}$$

and the model becomes self-similar [7–10].

In Fig. 1, we plot the super-diffusion behavior. Circles correspond to empirical data from S&P 500 cash index during the period January 1988–December 1996. The solid line shows the super-diffusive character predicted by Eq. (17) setting $\alpha = 1.30$ and $ab^{2/\alpha} = 2.44 \times 10^{-8}$ (if time is measured in minutes). Dashed line represents normal-diffusion $\langle X^2(t) \rangle \propto t$. Observe that data obey super-diffusion for $t \leqslant 10$ min, and when $t > 10$ min there seems to be a "crossover" to normal diffusion.

We finally study the asymptotic behavior of our distribution. It can be shown from Eq. (12) that the center of the distribution, defined by $|x| < (bt)^{1/\alpha}$, is again approximated by the Lévy distribution defined above. On the other hand, the tails of the distribution are solely determined by the jump pdf $h(u)$ by means of the expression

$$p(x, t) \sim \frac{abt}{|x|^{1+\alpha}} \int_{|x|/\sigma_m}^{\infty} u^\alpha h(u)\,\mathrm{d}u, \quad (|x| \gg (bt)^{1/\alpha}). \tag{20}$$

Therefore, return distributions present fat tails and have finite moments *if jump distributions behave in the same way*. This, in turn, allows us to make a statistical hypothesis on the form of $h(u)$ based on the empirical form and moments of the pdf.

In Fig. 2, we plot the probability density $p(x, t)$ of the S&P 500 cash index returns $X(t)$ observed at time $t = 1$ min (circles). $\Sigma = 1.87 \times 10^{-4}$ is the standard deviation of the empirical data. The dotted line corresponds to a Gaussian density with standard
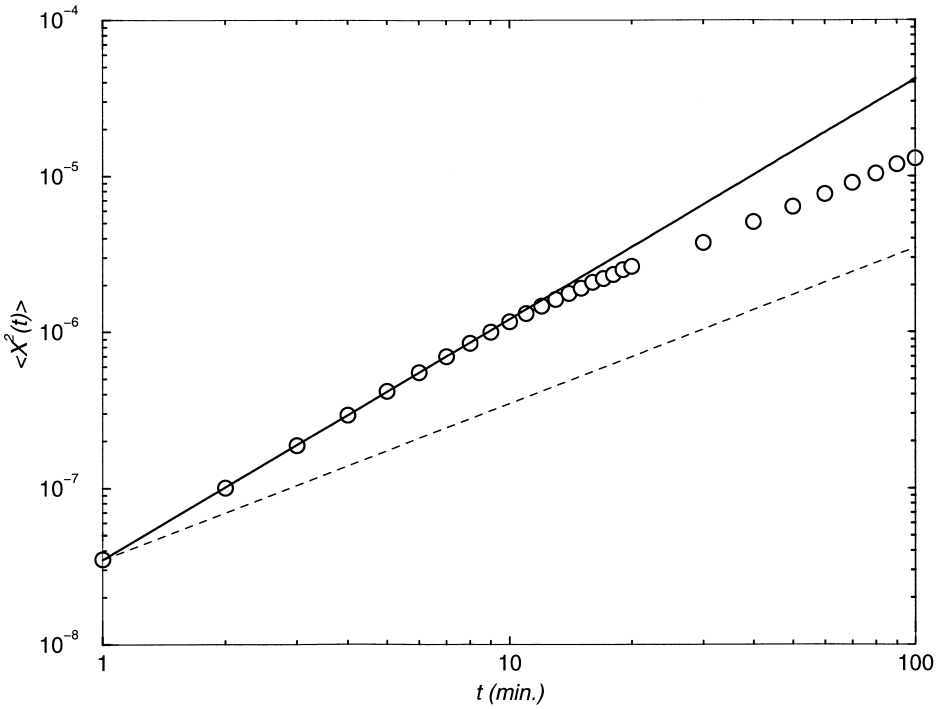
Fig. 1. Second moment of the zero-mean return. Circles correspond to empirical data from S&P 500 cash index (January 1988–December 1996). Solid line shows the super-diffusive character predicted by Eq. (17).

deviation given by Σ. The solid line shows the Fourier inversion of Eq. (12) with $\alpha = 1.30$, $\sigma_m = 9.07 \times 10^{-4}$, and $a = 2.97 \times 10^{-3}$. We use the gamma distribution of the absolute value of jump amplitudes,

$$h(u) = \mu^{\beta}|u|^{\beta-1}e^{-\mu|u|}/2\Gamma(\beta), \tag{21}$$

with $\beta = 2.39$, and $\mu = \sqrt{\beta(\beta+1)} = 2.85$. The dashed line represents a symmetrical Lévy stable distribution of index $\alpha = 1.30$ and the scale factor $k = 4.31 \times 10^{-6}$ obtained from Eq. (14). We note that the values of σ_m and Σ predict that the Pareto–Lévy distribution fails to be an accurate description of the empirical pdf for $x \gg 5\Sigma$ (see Eq. (20)).

We chose a gamma distribution of jumps because (i) as suggested by the empirical data analyzed, the tails of $p(x,t)$ decay exponentially, and (ii) one does not favor too small size jumps, i.e., those jumps with almost zero amplitudes. In any case, it would be very useful to obtain a more microscopic approach (based, for instance, in a "many agents" model [3,22]) giving some inside on the particular form of $h(u)$.
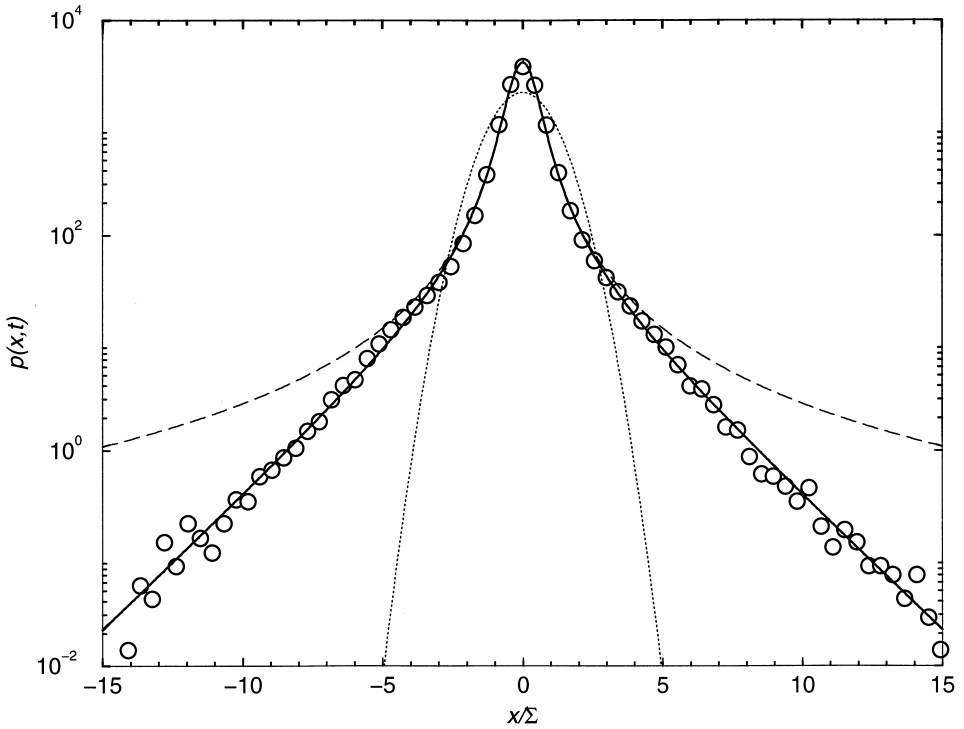
Fig. 2. Probability density function $p(x,t)$ for $t = 1$ min. Circles represent empirical data from S&P 500 cash index (January 1988–December 1996). Σ is the standard deviation of empirical data. The dotted line corresponds to the Gaussian density. The dashed line is the Lévy distribution and the solid line is the Fourier inversion of Eq. (22) with a gamma distribution of jumps (see the text for details).

4. Conclusions

Summarizing, by means of a continuous description of random pulses, we have obtained a dynamical model leading to a probability distribution for the speculative price changes. This distribution which is given by the following characteristic function:

$$\tilde{p}(\omega,t) = \exp\left\{ -a \int_0^1 z^{-1-\alpha}[1 - \tilde{h}(\omega z \sigma_m(t))]\,\mathrm{d}z \right\}, \tag{22}$$

where $\sigma_m(t) = (bt)^{1/\alpha}$, depends on three positive constants: $a, b,$ and $\alpha < 2$. The characteristic function (22) also depends on an unknown function $\tilde{h}(\omega)$, the unit-variance characteristic function of jumps, also to be conjectured and fitted from the tails of the empirical distribution. Therefore, starting from simple and reasonable assumptions we have developed a new stochastic process that possesses many of the features, i.e., fat tails, self-similarity, superdiffusion, and finite moments, of financial time series, thus providing us with a different point of view on the dynamics of the market. We finally point out that the model does not explain any correlation observed in empirical data

as some markets seem to have [4,23,24]. This insufficiency is due to the fact that we have modelled the behavior of returns through a mixture of independent sources of white noise. The extension of the model to include non-white noise sources and, hence, correlations will be presented soon.

Acknowledgements

References

[1] P.H. Cootner (Ed.), The Random Character of Stock Market Prices, MIT Press, Cambridge, MA, 1964.
[2] M.G. Kendall, J. Roy. Stat. Soc. 96 (1953) 11–25.
[3] J.Y. Campbell, A.W. Lo, A.C. MacKinlay, The Econometrics of Financial Markets, Princeton University Press, Princeton, NJ, 1997.
[4] J.P. Bouchaud, M. Potters, Théorie des Riskes Financiers, Aléa-Saclay, Paris, 1997.
[5] B. Mandelbrot, J. Business 35 (1963) 394–419.
[6] W. Feller, An Introduction to Probability Theory and its Applications, Wiley, New York, 1971.
[7] R.N. Mantegna, E.H. Stanley, Nature 376 (1995) 46–49.
[8] E. Scalas, Physica A 253 (1998) 394–402.
[9] S. Gashghaie, W. Breymann, J. Peinke, P. Talkner, Y. Dodge, Nature 381 (1996) 767–770.
[10] S. Gallucio, G. Caldarelli, M. Marsili, Y.-C. Zhang, Physica A 245 (1997) 423–436.
[11] E. Fama, J. Business 35 (1963) 420–429.
[12] R.N. Mantegna, H.E. Stanley, Phys. Rev. Lett. 73 (1994) 2946–2949.
[13] I. Koponen, Phys. Rev. E 52 (1995) 1197–1199.
[14] T. Bollerslev, R.Y. Chou, K.F. Kroner, J. Econometrics 52 (1992) 5–59.
[15] R.F. Engle, Econometrica 50 (1982) 987–1007.
[16] T. Bollerslev, J. Econometrics 31 (1986) 307–327.
[17] R.C. Merton, J. Financial Econom. 3 (1976) 125–144.
[18] J. Masoliver, Phys. Rev. A 35 (1987) 3918–3928.
[19] S.O. Rice, in: N. Wax (Ed.), Noise and Stochastic Processes, Dover, New York, 1954.
[20] R.N. Mantegna, H.E. Stanley, Nature 383 (1996) 587–588.
[21] R.N. Mantegna, H.E. Stanley, Physica A 239 (1997) 255–266.
[22] T. Lux, M. Marchesi, Nature 397 (1999) 498–500.
[23] A.W. Lo, A.C. MacKinlay, Rev. Financial Studies 1 (1988) 41–66.
[24] A.W. Lo, A.C. MacKinlay, Rev. Financial Studies 3 (1990) 175–205.