ELSEVIER

Modeling long-range memory trading activity by stochastic differential equations

V. Gontis*, B. Kaulakys

Institute of Theoretical Physics and Astronomy of Vilnius University, A. Goštauto 12, LT-01108 Vilnius, Lithuania

Available online 1 March 2007

Abstract

We propose a model of fractal point process driven by the nonlinear stochastic differential equation. The model is adjusted to the empirical data of trading activity in financial markets. This reproduces the probability distribution function and power spectral density of trading activity observed in the stock markets. We present a simple stochastic relation between the trading activity and return, which enables us to reproduce long-range memory statistical properties of volatility by numerical calculations based on the proposed fractal point process.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Stochastic equations; Point processes; Financial markets

1. Introduction

There are empirical evidences that the trading activity, the trading volume and the volatility of the financial markets are stochastic variables with the power-law probability distribution function (PDF) [1,2] and the long-range correlations [3–5]. Most of proposed models apply generic multiplicative noise responsible for the power-law PDF, whereas the long-range memory aspect is not accounted in the widespread models [6]. The additive–multiplicative stochastic models of the financial mean-reverting processes provide rich spectrum of shapes for PDF, depending on the model parameters [7], however, do not describe the long-memory features. Empirical analysis confirms that the long-range correlations in volatility arise due to those of trading activity [4]. On the other hand, trading activity is a financial variable dependant on the one stochastic process, i.e., interevent time between successive market trades. Therefore, it can be modeled as event flow of the stochastic point process.

Recently, we investigated analytically and numerically the properties of the stochastic multiplicative point processes, derived formula for the power spectrum [8,9] and related the model with the multiplicative stochastic differential equations [10]. Preliminary comparison of the model with the empirical data of the power spectrum and probability distribution of stock market trading activity [11] stimulated us to work on the more detailed definition of the model. Here we present the stochastic model of the trading activity with the long-range correlations and investigate its connection to the stochastic modeling of the volatility.

---

*Corresponding author.

E-mail address: gontis@itpa.lt (V. Gontis).

URL: http://www.itpa.lt/~gontis

The proposed stochastic nonlinear differential equations reproduce the power spectrum and PDF of the trading activity in the financial markets, describe the stochastic interevent time as the fractal-based point process and can be applicable for modeling of the volatility with the long-range autocorrelation.

2. Modeling fractal point process by the nonlinear stochastic differential equation

Trades in financial markets occur at discrete times $t_1, t_2, \ldots, t_k, \ldots$ and can be considered as identical point events. Such point process is stochastic and totally defined by the stochastic interevent time $\tau_k = t_{k+1} - t_k$. A fractal stochastic point process results, when at least two statistics exhibit the power-law scaling, indicating that represented phenomena contains clusters of events over all scales of time [12]. The dimension of the fractal point process is a measure of the clustering of the events within the process and by the definition coincides with the exponent of the power spectral density of the flow of events.

We can model the trading activity in financial markets by the fractal point process as its empirical PDF and the power spectral density exhibit the power-law scaling [4,13]. In this paper, we consider the flow of trades in financial markets as Poisson process driven by the multiplicative stochastic equation. First of all we define the stochastic rate $n = 1/\tau$ of event flow by continuous stochastic differential equation

$$d\tau = \gamma \tau^{2\mu-2} \, dt + \sigma \tau^{\mu-1/2} \, dW, \tag{1}$$

where W is a standard random Wiener process, σ denotes the standard deviation of the white noise, $\gamma \ll 1$ is a coefficient of the nonlinear damping and μ defines the power of noise multiplicativity. The diffusion of τ should be restricted at least from the side of high values. Therefore we introduce an additional term $-(m/2)\sigma^2(\tau/\tau_0)^m \tau^{2\mu-2}$ into Eq. (1), which produces the exponential diffusion reversion in equation

$$d\tau = \left[\gamma - \frac{m}{2}\sigma^2 \left(\frac{\tau}{\tau_0}\right)^m\right] \tau^{2\mu-2} \, dt + \sigma \tau^{\mu-1/2} \, dW, \tag{2}$$

where m and τ_0 are the power and value of the diffusion reversion, respectively. The associated Fokker–Planck equation with the zero flow gives the simple stationary PDF

$$P(\tau) \sim \tau^{\alpha+1} \exp\left[-\left(\frac{\tau}{\tau_0}\right)^m\right] \tag{3}$$

with $\alpha = 2(\gamma_\sigma - \mu)$ and $\gamma_\sigma = \gamma/\sigma^2$. Eq. (2) describes continuous stochastic variable τ, defines the rate $n = 1/\tau$ and, after the Ito transform of variable, results in stochastic differential equation

$$dn = \sigma^2 \left[(1 - \gamma_\sigma) + \frac{m}{2}\left(\frac{n_0}{n}\right)^m\right] n^{2\eta-1} \, dt + \sigma n^\eta \, dW, \tag{4}$$

where $\eta = \frac{5}{2} - \mu$ and $n_0 = 1/\tau_0$. Eq. (4) describes stochastic process n with PDF

$$P(n) \sim \frac{1}{n^\lambda} \exp\left\{-\left(\frac{n_0}{n}\right)^m\right\}, \quad \lambda = 2(\eta - 1 + \gamma_\sigma), \tag{5}$$

and power spectrum $S(f)$ [8–10]:

$$S(f) \sim \frac{1}{f^\beta}, \quad \beta = 2 - \frac{3 - 2\gamma_\sigma}{2\eta - 2}. \tag{6}$$

Noteworthy, that in the proposed model only two parameters, γ_σ and η (or μ), define exponents λ and β of two power-law statistics, i.e., of PDF and of the power spectrum. Time scaling parameter σ^2 in Eq. (4) can be omitted adjusting the time scale. Here we define the fractal point process driven by the stochastic differential equation (4) or equivalently by Eq. (2), i.e., we assume $\tau(t)$ as slowly diffusing mean interevent time of Poisson process with the stochastic rate n. This should produce the fractal point process with the statistical properties defined by Eqs. (5) and (6). Within this assumption the conditional probability of interevent time τ_p in the modulated Poisson point process with the stochastic rate $1/\tau$ is

$$\varphi(\tau_p|\tau) = \frac{1}{\tau} \exp\left[-\frac{\tau_p}{\tau}\right]. \tag{7}$$

Then the long-time distribution $P_k(\tau_p)$ of interevent time τ_p in k-space [8] has the integral form

$$P_k(\tau_p) = C \int_0^\infty \exp\left[-\frac{\tau_p}{\tau}\right] \tau^{\alpha-1} \exp\left[-\left(\frac{\tau}{\tau_0}\right)^m\right] d\tau, \tag{8}$$

with C defined from the normalization, $\int_0^\infty P_k(\tau_p)\,d\tau_p = 1$. In the case of pure exponential diffusion reversion, $m = 1$, PDF (8) has a simple form

$$P_k(\tau_p) = \frac{2}{\Gamma(1+\alpha)\tau_0}\left(\frac{\tau_p}{\tau_0}\right)^{\alpha/2} K_\alpha\left(2\sqrt{\frac{\tau_p}{\tau_0}}\right), \tag{9}$$

where $K_\alpha(z)$ denotes the modified Bessel function of the second kind. For $m > 1$ more complicated structures of distribution $P_k(\tau_p)$ expressed in terms of hypergeometric functions arise.

3. Adjustment of the model to the empirical data

We will investigate how the proposed modulated Poisson stochastic point process can be adjusted to the empirical trading activity, defined as number of transactions in the selected time window τ_d. Stochastic variable n denotes the number of events per unit time interval. One has to integrate the stochastic signal equation (4) in the time interval τ_d to get the number of events in the selected time window. In this paper we denote the integrated number of events as

$$N(t,\tau_d) = \int_t^{t+\tau_d} n(t')\,dt' \tag{10}$$

and call it the trading activity in the case of the financial market.

Detrended fluctuation analysis [13] is one of the methods to analyze the second order statistics related to the autocorrelation of trading activity. The exponents v of the detrended fluctuation analysis obtained by fits for each of the 1000 US stocks show a relatively narrow spread of v around the mean value $v = 0.85 \pm 0.01$ [13]. We use relation $\beta = 2v - 1$ between the exponents v of detrended fluctuation analysis and the exponents β of the power spectrum [14] and in this way define the empirical value of the exponent for the power spectral density $\beta = 0.7$.

Our analysis of the Lithuanian stock exchange data confirmed that the power spectrum of trading activity is the same for various liquid stocks even for the emerging markets [11]. The histogram of exponents obtained by fits to the cumulative distributions of trading activities of 1000 US stocks [13] gives the value of exponent $\lambda = 4.4 \pm 0.05$ describing the power-law behavior of the trading activity. Empirical values of $\beta = 0.7$ and $\lambda = 4.4$ confirm that the time series of the trading activity in real markets are fractal with the power-law statistics. Time series generated by stochastic process (4) are fractal in the same sense.

Nevertheless, we face serious complications trying to adjust model parameters to the empirical data of the financial markets. For the pure multiplicative model, when $\mu = 1$ or $\eta = \frac{3}{2}$, we have to take $\gamma_\sigma = 0.85$ to get $\beta = 0.7$ and $\gamma_\sigma = 1.7$ to get $\lambda = 4.4$, i.e., it is impossible to reproduce the empirical PDF and power spectrum with the same relaxation parameter γ_σ and exponent of multiplicativity μ. We have proposed possible solution of this problem in our previous publications [9,11] deriving PDF for the trading activity N:

$$P(N) \sim \begin{cases} \dfrac{1}{N^{3+\alpha}}, & N \ll \gamma^{-1}, \\[2mm] \dfrac{1}{N^{5+2\alpha}}, & N \gg \gamma^{-1}. \end{cases} \tag{11}$$

When $N \gg \gamma^{-1}$ this yields exactly the required value of $\lambda = 5 + 2\alpha = 4.4$ and $\beta = 0.7$ for $\gamma_\sigma = 0.85$.

Nevertheless, we cannot accept this as the sufficiently accurate model of the trading activity since the empirical power law distribution is achieved only for very high values of the trading activity. Probably this reveals the mechanism how the power-law distribution converges to normal distribution through the growing values of the exponent, but empirically observed power-law distribution in wide area of N values cannot be reproduced. Let us notice here that the desirable power-law distribution of the trading activity with the

exponent $\lambda = 4.4$ may be generated by the model (4) with $\eta = \frac{5}{2}$ and $\gamma_\sigma = 0.7$. Moreover, only the smallest values of τ or high values of n contribute to the power spectral density of trading activity [10]. This suggests us to combine the stochastic process with two values of μ: (i) $\mu \simeq 0$ for the main area of τ and n diffusion and (ii) $\mu = 1$ for the lowest values of τ or highest values of n. Therefore, we introduce a new stochastic differential equation for n combining two powers of the multiplicative noise,

$$dn = \sigma^2 \left[(1 - \gamma_\sigma) + \frac{m}{2} \left(\frac{n_0}{n} \right)^m \right] \frac{n^4}{(n\varepsilon + 1)^2} \, dt + \frac{\sigma n^{5/2}}{(n\varepsilon + 1)} \, dW, \tag{12}$$

where a new parameter ε defines crossover between two areas of n diffusion. The corresponding iterative equation for τ_k in such a case is

$$\tau_{k+1} = \tau_k + \left[\gamma - \frac{m}{2} \sigma^2 \left(\frac{\tau}{\tau_0} \right)^m \right] \frac{\tau_k}{(\varepsilon + \tau_k)^2} + \sigma \frac{\tau_k}{\varepsilon + \tau_k} \varepsilon_k, \tag{13}$$

where ε_k denotes uncorrelated normally distributed random variable with the zero expectation and unit variance.

Eqs. (12) and (13) define related stochastic variables n and τ, respectively, and they should reproduce the long-range statistical properties of the trading activity and of waiting time in the financial markets. We verify this by the numerical calculations. In Fig. 1 we present the power spectral density calculated for the equivalent processes (12) and (13) (see Ref. [9] for details of calculations). This approach reveals the structure of the power spectral density in wide range of frequencies and shows that the model exhibits not one but rather two separate power laws with the exponents $\beta_1 = 0.33$ and $\beta_2 = 0.72$. From many numerical calculations performed with the multiplicative point processes we can conclude that combination of two power laws of spectral density arises only when the multiplicative noise is a crossover of two power laws as in Eqs. (12) and (13). Let us choose the value of crossover parameter $\varepsilon = 0.07$ to balance the influence of two rather different market behaviors. We will show in the next section that this may serve as an explanation of two exponents of the power spectrum in the empirical data of volatility for S&P 500 companies [15].

Empirical data of the trading activity statistics should be modeled by the integrated flow of events N defined in the time interval $\tau_d \geqslant \tau_0$. In Fig. 2 we demonstrate the probability distribution functions $P(N)$ and its cumulative form $P_>(N)$ calculated from the histogram of N generated by Eq. (13) with the selected time interval $\tau_d = 100$. This illustrates that the model distribution of the integrated signal N has the power-law form with the same exponent $\lambda = 4.4$ as observed in empirical data [4,5].

The power spectrum of the trading activity N has the same exponent $\beta = 0.7$ as power spectrum of n in the low frequency area for all values of τ_d.
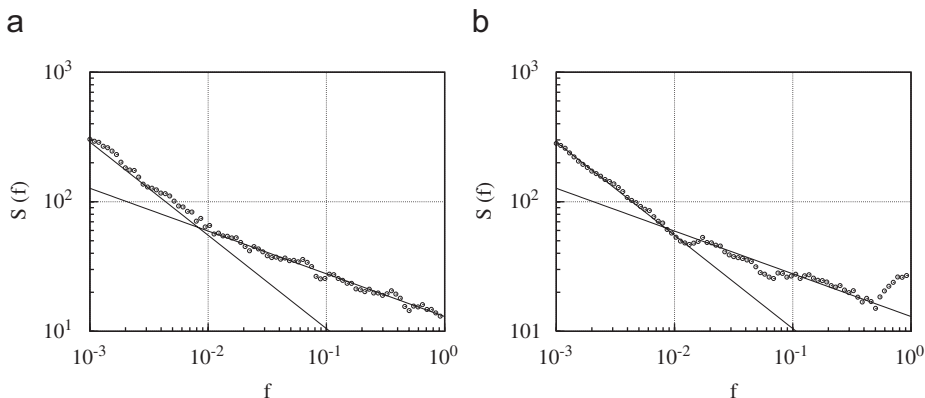

Fig. 1. Power spectral density $S(f)$ averaged over 100 realizations of series with 1 000 000 iterations and parameters $\gamma = 0.0004$; $\sigma = 0.025$; $\varepsilon = 0.07$; $\tau_0 = 1$; $m = 6$. Straight lines approximate power spectrum $S \sim 1/f^{\beta_{1,2}}$ with $\beta_1 = 0.33$ and $\beta_2 = 0.72$: (a) $S(f)$ of the flow $I(t) = \sum_k \delta(t - t_k)$ with the interevent time $\tau_k = t_{k+1} - t_k$ generated by Eq. (13), (b) $S(f)$ calculated by the Fast Fourier Transform of n series generated by Eq. (12).
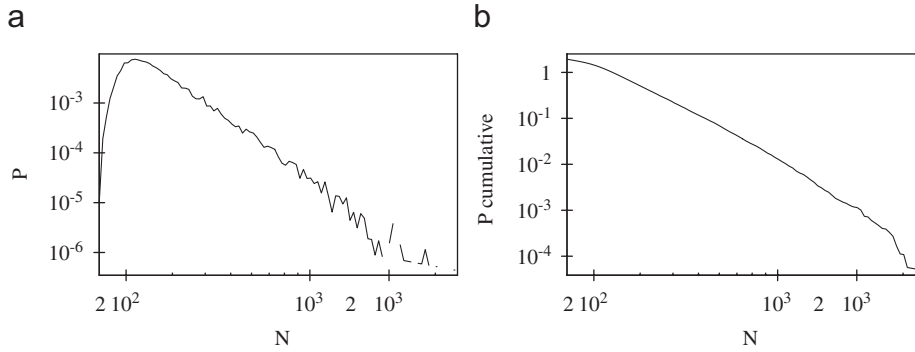
Fig. 2. (a) PDF $P(N)$ calculated from the histogram of N generated by Eq. (13) with the selected time interval $\tau_d = 100$. (b) Cumulative distribution $P_>(N)$. Other parameters are as in Fig. 1.
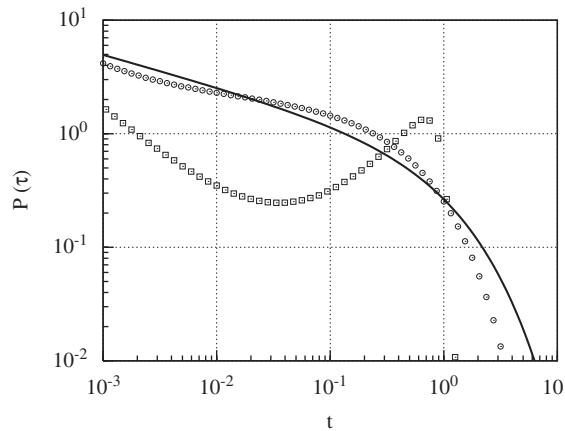

Fig. 3. PDF of interevent time $P(\tau)$: open circles, calculated from the histogram of τ_p generated by Eq. (7) with the mean interevent time calculated from Eq. (14); open squares, calculated from the iterative equation (13). Used parameters are as in Fig. 1. Solid line represents the scaling function (3) of Ref. [16] with $\delta = 0.72$, adapted to the empirical data of Fig. 2(c) in Ref. [16].

The same numerical results can be reproduced by continuous stochastic differential equation (12) or iteration equation (13). One can consider the discrete iterative equation for the interevent time τ_k (13) as a method to solve numerically continuous equation

$$\mathrm{d}\tau = \left[\gamma - \frac{m}{2}\sigma^2\left(\frac{\tau}{\tau_0}\right)^m\right]\frac{1}{(\varepsilon + \tau)^2}\,\mathrm{d}t + \sigma\frac{\sqrt{\tau}}{\varepsilon + \tau}\,\mathrm{d}W. \tag{14}$$

The continuous equation (12) follows from the Eq. (14) after change of variables $n = 1/\tau$.

We can conclude that the long-range memory properties of the trading activity in the financial markets as well as the PDF can be modeled by the continuous stochastic differential equation (12). In this model the exponents of the power spectral density, β, and of PDF, λ, are defined by one parameter $\gamma_\sigma = \gamma/\sigma^2$. We consider the continuous equation of the mean interevent time τ as a model of slowly varying stochastic rate $1/\tau$ in the modulated Poisson process (7). In Fig. 3, we demonstrate the probability distribution functions $P(\tau_p)$ calculated from the histogram of τ_p generated by Eq. (7) with the diffusing mean interevent time calculated from Eq. (14), open circles. The curve of open squares demonstrates different areas of τ diffusion depending on crossover parameter $\varepsilon = 0.07$. In the modulated Poisson process, represented by open circles, and the empirical data, represented by solid curve, the histograms are flattened by Poisson nature of the process.

Numerical results show good qualitative agreement with the empirical data of interevent time probability distribution measured from few years series of U.S. stock data [16]. This enables us to conclude that the proposed stochastic model captures the main statistical properties including PDF and the long-range correlation of the trading activity in the financial markets. Furthermore, in the next section we will show that this may serve as a background statistical model responsible for the statistics of return volatility in widely accepted geometric Brownian motion (GBM) of the financial asset prices.

4. Modeling long-range memory volatility

The basic quantities studied for the individual stocks are price $p(t)$ and return

$$x(t, \tau_{\mathrm{d}}) = \ln p(t + \tau_{\mathrm{d}}) - \ln p(t). \tag{15}$$

Let us express return $x(t, \tau_{\mathrm{d}})$ over a time interval τ_{d} through the subsequent changes δx_i due to the trades $i = 1, 2, \ldots, N(t, \tau_{\mathrm{d}})$ in the time interval $[t, t + \tau_{\mathrm{d}}]$,

$$x(t, \tau_{\mathrm{d}}) = \sum_{i=1}^{N(t, \tau_{\mathrm{d}})} \delta x_i. \tag{16}$$

We denote the variance of δx_i calculated over the time interval τ_{d} as $W^2(t, \tau_{\mathrm{d}})$. If δx_i are mutually independent one can apply the central limit theorem to sum (16). This implies that for the fixed variance $W^2(t, \tau_{\mathrm{d}})$ return $x(t, \tau_{\mathrm{d}})$ is a normally distributed random variable with the variance $W^2(t, \tau_{\mathrm{d}}) N(t, \tau_{\mathrm{d}})$

$$x(t, \tau_{\mathrm{d}}) = W(t, \tau_{\mathrm{d}}) \sqrt{N(t, \tau_{\mathrm{d}})} \varepsilon_t, \tag{17}$$

where ε_t is the normally distributed random variable with the zero expectation and unit variance.

Empirical test of conditional probability $P(x(t, \tau_{\mathrm{d}}) | W(t, \tau_{\mathrm{d}}))$ [4] confirms its Gaussian form, and the unconditional distribution $P(x(t, \tau_{\mathrm{d}}))$ is a power law with the cumulative exponent 3. This implies that the power-law tails of returns are largely due to those of $W(t, \tau_{\mathrm{d}})$. Here we refer to the theory of price diffusion as a mechanistic random process [17,18]. For this idealized model the short-term price diffusion depends on the limit order removal and this way is related to the market order flow. Furthermore, the empirical analysis confirms that the volatility calculated for the fixed number of transactions has the long memory properties as well and it is correlated with real time volatility [19]. We accumulate all these results into the assumption that standard deviation $W(t, \tau_{\mathrm{d}})$ may be proportional to the square root of the trading activity, i.e., $W(t, \tau_{\mathrm{d}}) \sim k \sqrt{N(t, \tau_{\mathrm{d}})}$. This enables us to propose a simple model of return

$$x(t, \tau_{\mathrm{d}}) \sim k N(t, \tau_{\mathrm{d}}) \varepsilon_t \tag{18}$$

and related model of volatility $v = |x(t, \tau_{\mathrm{d}})|$ based on the proposed model of trading activity (12). We generate series of trade flow $n(t)$ numerically solving Eq. (12) with variable steps of time $\Delta t_i = h_i = n_0/n_i$ and calculate
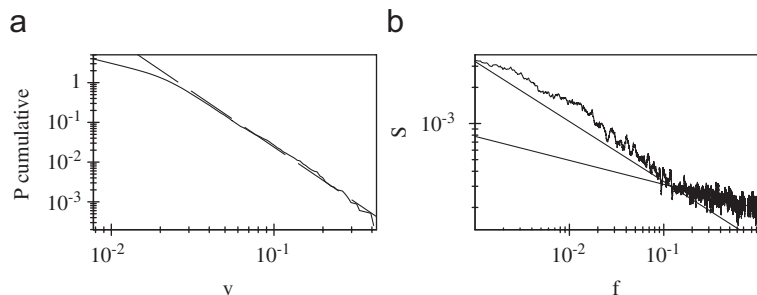

Fig. 4. (a) Cumulative probability distribution function of the volatility, $P_>(\bar{v})$, averaged over 10 intervals calculated from the series of $n(t)$ generated by Eqs. (12) and (18), all parameters are the same as in previous calculations. Dashed line approximates the power law $P(\bar{v}) \sim 1/\bar{v}^{2.8}$. (b) Power spectral density $S(f)$ of v calculated from FFT of the same series $n(t)$. Straight lines approximate the empirical power spectral density $S \sim 1/f^{\beta_{1,2}}$ with $\beta_1 = 0.5$ and $\beta_2 = 0.2$ presented in Fig. 10a of Ref. [15].

the trading activity in subsequent time intervals τ_d as $N(t, \tau_d) = \int_t^{t+\tau_d} n(t')\,dt'$. This enables us to generate series of return $x(t, \tau_d)$, of volatility $v(t, \tau_d) = |x(t, \tau_d)|$ and of the averaged volatility $\bar{v} = (1/m)\sum_{i=1}^{i=m} v(t_i, \tau_d)$.

In Fig. 4 we demonstrate cumulative distribution of $\bar{v}$ and the power spectral density of $v(t, \tau_d)$ calculated from FFT. We see that proposed model enables us to catch up the main features of the volatility: the power-law distribution with exponent 2.8 and power spectral density with two exponents $\beta_1 = 0.5$ and $\beta_2 = 0.2$. This is in a good agreement with the empirical data [15,19].

5. Conclusions

Starting from the concept of the fractal point processes [12] we proposed process driven by the nonlinear stochastic differential equation and based on the earlier introduced stochastic point process model [8–11]. This may serve as a possible model of the flow of points or events in the physical, biological and social systems when their statistics exhibit power-law scaling indicating that the represented phenomena contains clusters of events over all scales. First of all, we analyze the statistical properties of trading activity and waiting time in financial markets by the proposed Poisson process with the stochastic rate defined as a stand-alone stochastic variable. We consider the stochastic rate as continuous one and model it by the stochastic differential equation, exhibiting long-range memory properties [10]. Further we propose a new form of the stochastic differential equation combining two powers of multiplicative noise: one responsible for the probability distribution function and another responsible for the power spectral density. The proposed new form of the continuous stochastic differential equation enabled us to reproduce the main statistical properties of the trading activity and waiting time, observable in the financial markets. In the new model the power spectral density with two different scaling exponents arise. This is in agreement with the empirical power spectrum of volatility and implies that the market behavior may be dependant on the level of activity. One can observe at least two stages in market behavior: calm and excited. Finally, we propose a very simple stochastic relation between trading activity and return to reproduce the statistical properties of volatility. This enabled us to model empirical distribution and long-range memory of volatility.

Acknowledgments

References

[1] B.B. Mandelbrot, J. Business 36 (1963) 394.
[2] T. Lux, Appl. Finance Econ. 6 (1996) 463.
[3] R.F. Engle, A.J. Patton, Quant. Finance 1 (2001) 237.
[4] V. Plerou, P. Gopikrishnan, X. Gabaix, L.A.N. Amaral, H.E. Stanley, Quant. Finance 1 (2001) 262.
[5] X. Gabaix, P. Gopikrishnan, V. Plerou, H.E. Stanley, Nature 423 (2003) 267.
[6] M.M. Dacorogna, R. Gencay, U.A. Muller, R.B. Olsen, O.V. Pictet, An Introduction to High-Frequency Finance, Academic Press, San Diego, 2001.
[7] C. Anteneodo, R. Riera, Phys. Rev. E 72 (2005) 026106.
[8] B. Kaulakys, V. Gontis, M. Alaburda, Phys. Rev. E 71 (2005) 051105.
[9] V. Gontis, B. Kaulakys, Physica A 343 (2004) 505.
[10] B. Kaulakys, J. Ruseckas, V. Gontis, M. Alaburda, Physica A 365 (2006) 217.
[11] V. Gontis, B. Kaulakys, Physica A 344 (2004) 128.
[12] S.B. Lowen, M.C. Teich, Fractal-Based Point Processes, Wiley, New York, 2005 (ISBN: 0-471-38376-7).
[13] V.E. Plerou, P. Gopikrishnan, L. Amaral, X. Gabaix, E. Stanley, Phys. Rev. E 62 (2000) R3023.
[14] J. Beran, Statistics for Long-Memory Processes, Chapman & Hall, NY, 1994.
[15] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, Ch. Peng, H.E. Stanley, Phys. Rev. E 60 (1999) 1390.
[16] P. Ivanov, A. Yuen, B. Podobnik, Y. Lee, Phys. Rev. E 69 (2004) 056107.
[17] M. Daniels, D. Farmer, L. Gillemot, G. Iori, E. Smith, Phys. Rev. Lett. 90 (2003) 108102.
[18] D. Farmer, L. Gillemot, F. Lillo, M. Szabolcs, A. Sen, Quant. Finance 4 (2004) 383.
[19] L. Gillemot, J.D. Farmer, F. Lillo, Santa Fe Institute Working Paper 05-12-041.