Journal of Statistical Physics, Vol. 89, Nos. 1/2, 1997

Econophysics: Scaling and Its Breakdown in Finance

Rosario N. Mantegna[1] and H. Eugene Stanley[2]

Received February 17, 1997; final March 16, 1997

We discuss recent empirical results obtained by analyzing high-frequency data of a stock market index, the Standard and Poor's 500. We focus on the scaling properties and on its breakdown of the index dynamics. A simple stochastic model, the truncated Lévy flight, is illustrated. Successes and limitations of this model are presented. A discussion about similarities and differences between the scaling properties observed in financial markets and in fully developed turbulence is also provided.

KEY WORDS: Random walks; Lévy stable processes; scaling; turbulence.

1. INTRODUCTION

Econophysics is an interdisciplinary subfield, with a growing number of practitioners. We shall briefly describe the spirit and substance of some recent work that focuses on scaling and its breakdown in financial data.

During the last thirty years, physicists have achieved important results in the field of phase transitions, statistical mechanics, nonlinear dynamics disordered and self-organized systems. In these fields, power laws, scaling and unpredictable (stochastic or deterministic) time series are present and the current interpretation of the underlying physics is often obtained using these concepts.

With this background in mind, it may surprise scholars trained in the natural science to learn that one of the first use of a power-law distribution took place in the social sciences. Almost exactly 100 years ago, the Italian social economist Vilfredo Pareto investigated the statistical character of

---

Dedicated to Bernard Jancovici on the occasion of his 65th birthday.

[1] Istituto Nazionale di Fisica della Materia, Unità di Palermo and Dipartimento di Energetica ed Applicazioni di Fisica, Università di Palermo, Palermo, I-90128, Italy.

[2] Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215.

0022-4715/97/1000-0469$12.50/0

the wealth of individuals in a stable economy by modeling them using the distribution

$$y = x^{-\nu}, \tag{1}$$

where y is the number of people having income x or greater than x and ν is an exponent that Pareto estimated to be ≈ 1.5.[1] Pareto noticed that his result was quite general and applicable to nations "as different as those of England, of Ireland, of Germany, of the Italian cities, and even of Peru."

It should be fully appreciated that the concept of a power-law distribution is counter-intuitive. A power-law distribution lacks any characteristic scale. This property prevented the use of power-law distributions in the natural sciences until the recent emergence of new paradigms (i) in probability theory, thanks to the work of Paul Lévy and Kolmogorov and thanks to the application of power-law distributions to several problems pursued by Mandelbrot; and (ii) in the study of phase transitions in which the concept of scaling of thermodynamic potentials and correlation functions was introduced.

Other processes and concepts widely recurrent in statistical physics have been used quite early in the study of financial systems. Self-similarity in the price change distributions is implicit in the pioneering work of Bachelier.[2] Scaling (in the sense of power-law tails in the distribution of logarithmic changes of prices) and self-similar non-Gaussian distributions of logarithmic changes of prices were first proposed and tested by Mandelbrot in '63 by modeling and analyzing the statistical properties of cotton prices.[3]

In this paper, we briefly recall some recent empirical results obtained by analyzing the time evolution of the Standard & Poor 500 index of the New York Stock Exchange recorded with high temporal resolution. We also discuss the properties of a simple stochastic model, the truncated Lévy flight (TLF). This model is able to describe several of the major features observed in empirical data. Limitations of this simple model are also addressed. We end by discussing similarity and differences observed in the scaling properties of the price dynamics in a financial market and the dynamics of the velocity of a 3-dimensional turbulent fluid.

II. SCALING AND ITS BREAKDOWN IN THE S&P 500 INDEX

Stock exchange time series have been modelled as stochastic processes with very short time memory since the seminal study of Bachelier published at the beginning of this century.[2] Since then several stochastic models have been proposed and tested in the economics[3–9] and physics[10–14] literature to cite only a few studies. Alternative approaches based on the paradigm of chaotic dynamics have been also proposed.[15–17] The most widely-accepted models state that the variations of share price is a random process with very short time memory. For the distribution of variations of the logarithm of asset prices several proposal have been published. These include (i) a normal distribution,[2] (ii) a Lévy stable distribution,[3] (iii) leptokurtik distributions generated by a mixture of normal distributions[6] and (iv) ARCH/GARCH models.[7, 8]

The proposals of: (i) a normal distribution,[2] and (ii) a Lévy stable distribution[3] obey respectively the central-limit-theorem[18] or a generalized version of it.[19] The most striking difference between these two stochastic processes involves the wings of the distributions. Distinguishing between the two processes (i) and (ii) by comparing the distribution wings can be quite difficult because financial data sets (as all other data sets) are unavoidably limited. To maximize the amount of data to be analyzed in a limited time interval (limited to avoid that underlying rules or deep economics changes could happen inside the investigated period), we chose to investigate high frequency data.

Data, kindly provided by the Chicago Mercantile Exchange, consist of all 1,447,514 records of the S&P 500 cash index recorded during the 6-year period 1/84–12/89. The time intervals between successive records are not fixed: the average value between successive records is close to 1 min during 1984 and 1985 and close to 15 s during 1986–1989. We define the trading time as a continuous time starting from the opening of the day until the closing, and then continuing with the opening of the next trading day. From this data base, we select the complete set of non-overlapping records separated by a time interval $\Delta t \pm \varepsilon \Delta t$ (where ε is the tolerance, always less than 0.035). We denote the value of the S&P 500 as $y(t)$, and we define

$$z(t) \equiv y(t) - y(t - \Delta t). \tag{2a}$$

In the high frequency regime, the two stochastic processes $z(t)$ (price differences) and

$$r(t) \equiv \ln(y(t + \Delta t) - \ln(y(t)) \tag{2b}$$

logarithmic differences have similar statistical properties. This is due to the fact that in the regime $z(t) \ll y(t)$, $r(t) = \ln[1 + z(t)/y(t)]$ is bounded by

$$\frac{z(t)/y(t)}{1 + z(t)/y(t)} < r(t) < \frac{z(t)}{y(t)} \tag{3}$$

and $z(t)$ is a "fast" variable, whereas $y(t)$ is a "slow" variable being the integral of $z(t)$.

To quantitatively characterize the experimentally-observed process, we determine[14] the probability distribution $P(z)$ of index variations for different values of Δt. We select Δt values that are logarithmically equally spaced ranging from 1 to 1000 min. The number of data in each set is decreasing from the maximum value of 493,545 ($\Delta t = 1$ min) to the minimum value of 562 ($\Delta t = 1000$ min). We note[14] that the distributions are non-Gaussian, indeed, they have wings larger than expected for a normal process. A determination of the parameters characterizing the distributions is difficult if one uses methods that mainly investigate the wings of distributions, especially because larger values of Δt imply a reduced number of data.

Therefore we use a different approach: we study the "probability of return to the origin" $P(z=0)$ as a function of Δt. With this choice we investigate the point of each probability distribution that is least affected by the noise introduced by the finiteness of the experimental data set. Our investigation of $P(0)$ versus Δt in a log-log plot[14] shows that the data are well-fit by a straight line characterized by the slope -0.712 ± 0.025. We observe a non-normal scaling behavior (slope $\neq -0.5$) in an interval of trading time ranging from 1 to 1000 min.

This empirical finding agrees with the theoretical model of a Lévy flight[20] or Lévy walk.[21] In fact, if the central region of the distribution is well-described by a Lévy stable symmetrical distribution,[22]

$$L_\alpha(z, \Delta t) \equiv \frac{1}{\pi} \int_0^\infty \exp(-\gamma \Delta t q^\alpha) \cos(qz)\, dq, \tag{4}$$

of index α and scale factor γ at $\Delta t = 1$, then the probability of return is given by

$$P(0) \equiv L_\alpha(0, \Delta t) = \frac{\Gamma(1/\alpha)}{\pi\alpha(\gamma \Delta t)^{1/\alpha}}. \tag{5}$$

By using the value -0.712 from the analysis of the probability of return one obtains the index $\alpha = 1.40 \pm 0.05$.[14]

It is known from the literature[23] that the intraday variance of the logarithmic price or price difference is not constant. We perform a check to determine if this fact affects the scaling properties of the price difference. Specifically, we select from our database the price differences recorded during a trading day within the window time starting 90 minutes after the opening and ending 2 hours before closing. In this way, we do not take into account the overnight returns in our analysis and moreover we also do not use the periods of higher volatility known to occur at the beginning and end of each trading day. With this analysis we obtain $\alpha = -0.725 \pm 0.025$ a value very close to the one obtained for the complete set of data. Hence this test shows that the overnight returns do not affect significantly the scaling properties (α) of the price differences.

We also check if the scaling extends over the entire probability distribution as well as $z = 0$. All the distributions agree well with a Lévy stable distribution.[14, 24] The distributions obtained with the highest temporal resolution ($\Delta t < 10$) show that in addition to the good agreement with the Lévy (non-Gaussian) profile observed for almost three orders of magnitude an approximately exponential fall-off is present. The clear deviation of the tails of the distribution from the Lévy profile shows us that the experimental tails are less fat than expected for a Lévy distribution. The deviation from the Lévy distribution is not observable for $\Delta t \gtrsim 10$ due to the limited number of records used to obtain these distributions.

The Lévy distribution has an infinite second moment (if $\alpha < 2$).[22] However, our empirical finding of an exponential (or stretched exponential) fall-off implies that the second moment is finite, thereby resolving the question about the finiteness of the variance of the price change distribution.[25] This conclusion might at first sight seem to contradict our observation of Lévy scaling of the central part of the price difference distribution over fully three orders of magnitude. However, there is no contradiction since, for example, the above findings might be interpreted in terms of a simple stochastic process, the truncated Lévy flight.[26]

III. THE TRUNCATED LÉVY FLIGHT

The truncated Lévy flight (TLF) has been introduced by Mantegna and Stanley in ref. 26. A TLF is defined as a stochastic process $\{x\}$ characterized by the following probability density function

$$T(x) \equiv \begin{cases} 0 & x > \ell \\ c_1 L(x) & -\ell \leqslant x \leqslant \ell \\ 0 & x < -\ell \end{cases} \tag{6}$$

where

$$L(x) \equiv \frac{1}{\pi} \int_0^{+\infty} \exp(-\gamma q^\alpha) \cos(qx)\, dq \tag{7}$$

is the symmetrical Lévy stable distribution of index $\alpha (0 < \alpha \leqslant 2)$ and scale factor γ ($\gamma > 0$), c_1 is a normalizing constant and ℓ is the cutoff length. For the sake of simplicity, we set $\gamma = 1$.

The central limit theorem (CLT) is fundamental to statistical mechanics. It states that when $n \to \infty$, the sum

$$z_n \equiv \sum_{i=1}^{n} x_i \tag{8}$$

of n stochastic variables $\{x\}$ that are statistically independent, identically distributed and with a finite variance, converges to a normal (Gaussian) stochastic process. Generally, $n \approx 10$ is sufficient to ensure convergence. In a dynamical system, Eq. (8) defines a random walk if the variable x is the jump size performed after a time interval Δt and n is the number of time intervals. In this lecture, the "number of variables" n and the "time" $t = n\Delta t$ can be interchanged everywhere.

We investigate the probability distribution $P(z_n)$ of the stochastic process of Eq. (8) when $\{x\}$ is a TLF, i.e., a stochastic process with probability distribution given by Eq. (6). We monitor the degree of convergence of the TLF to the asymptotic normal process by investigating the probability of return to the origin of the process $P(z_n = 0)$. The reason for this choice is twofold, first this will give us a concrete parallel to what we investigated in the previous section, and second the point $z_n = 0$ of the distribution $P(z)$ is the last point to converge to the asymptotic normal process for symmetrical stochastic processes.

For low values of n, $P(z_n = 0)$ takes a value very close to the one expected for a Lévy stable process

$$P(z_n = 0) \simeq L(z_n = 0) = \frac{\Gamma(1/\alpha)}{\pi \alpha n^{1/\alpha}}. \tag{9}$$

For large values of n, $P(z_n = 0)$ assumes the value predicted for a normal process,

$$P(z_n = 0) \simeq N(z_n = 0) = \frac{1}{\sqrt{2\pi}\, \sigma_o(\alpha, \ell)\, n^{1/2}}, \tag{10}$$

where $\sigma_o(\alpha, \ell)$ is the standard deviation of the TLF stochastic process $\{x\}$.

In the interval $1 \leqslant \alpha < 2$, the crossover between the two regimes has been determined in ref. 26 as:

$$n_\times \simeq A \ell^\alpha, \tag{11}$$

where

$$A = \left[\frac{\pi\alpha}{2\Gamma(1/\alpha)[\Gamma(1+\alpha)\sin(\pi\alpha/2)/(2-\alpha)]^{1/2}} \right]^{2\alpha/(\alpha-2)} \tag{12}$$

The description of the convergence process is not crucially depending on the exact shape of the cut-off[27] and some results of ref. 26 have been confirmed analytically for an exponential cut-off in ref. 28.

By performing numerical simulations, it is possible to investigate the process of convergence of the TLF to its asymptotic Gaussian process. To generate a Lévy stable stochastic process of index α and scale factor $\gamma = 1$, we use the algorithm of ref. 29. Other algorithms can be found in the mathematical literature.[30] We verify that the probability of return to the origin indicates with high accuracy the degree of convergence of the process to one of the two asymptotic regimes.

To summarize, our study shows that by investigating the probability of return to the origin of an originally quasi-stable non-normal stochastic process with finite variance a clear crossover between Lévy and Gaussian regimes is observable. Hence a Lévy-like probability distribution can be experimentally observed for a long (but finite) interval of time (or number of variables) even in the presence of stochastic processes characterized by a finite variance. However, not all the features observed in the S&P 500 dynamics are described by the TLF model. The simplest version of the model cannot describe the short time memory (of the order of 20 minutes or less) observed in the empirical data[24, 31] and also does not explain the empirical observation of the time dependence of the parameter γ which is fluctuating with burst of activity localized in specific months.[14, 24] The γ parameter is related to what is called "volatility" in the economic literature.[32]

IV. ANALOGIES AND DIFFERENCES WITH TURBULENCE

Other models might also be considered to fully describe the stock market data. For example, by using a rather different approach, an alternative possible physical phenomenon that it is worthwhile to investigate is turbulence.[33] The goal is to see if turbulence might be used as a paradigm to describe some of the phenomena empirically observed in the analysis of data of the S&P 500 dynamics. The rational for this choice is that it is known that intermittency of the dissipation rate and non-Gaussian profile of the probability density function of velocity changes are observed in the time evolution of a fully turbulent fluid moving in a 3-dimensional space.

This research has been performed independently by different groups[31, 34, 35] using different data bases.

To investigate analogies and differences between the quantitative measures of fluctuations in an economic index and the fluctuations in velocity of a fluid in a fully turbulent state we have systematically compared[31] the statistical properties of the S&P 500 cash index with the statistical properties of the velocity of turbulent air.

The turbulence data were kindly provided by Prof. K. R. Sreenivasan. Measurements were made[36] in the atmospheric surface layer about 6 m above a wheat canopy in the Connecticut Agricultural research station. The Taylor microscale Reynolds number R_λ was of the order of 1500. The file consists of 130,000 velocity records $v(t)$ digitized and linearized before processing. The associated velocity differences is defined as $U_{\Delta t}(t) \equiv v(t) - v(t - \Delta t)$.

Quantitative parallel analysis have been performed[31] by measuring the time dependence of the standard deviations $\sigma_Z(\Delta t)$ and $\sigma_U(\Delta t)$ of $P(Z)$ and $P(U)$, we find that: (i) In the case of the S&P 500 index variations the time dependence of the standard deviation, when $\Delta t \geqslant 15$ minutes fits well the behavior

$$\sigma_Z(\Delta t) \propto (\Delta t)^{0.53}. \tag{13}$$

The exponent is close to the typical value of 0.5 observed in random processes with independent increments. (ii) The velocity difference of the fully turbulent fluid shows a time dependence of the standard deviation, fitting the behavior

$$\sigma_U(\Delta t) \propto (\Delta t)^{0.33} \tag{14}$$

which is observed in short-time anticorrelated random processes.

Equivalent conclusions are reached if we measure the spectral density of the time series $y(t)$ and $v(t)$.

Another difference between the two processes is observed by investigating the probability of return to the origin $P(U=0)$ as functions of the time interval Δt between successive observations. The deviation from a Gaussian process is measured by comparing $P(U=0)$ with the value of $P_g(0)$. $P_g(0)$ is determined starting from the measured values of $\sigma(\Delta t)$ by using the equation

$$P_g(0) = \frac{1}{\sqrt{2\pi}\,\sigma(\Delta t)} \tag{15}$$

valid for a Gaussian process.

We observe[31] a clear difference between $P(0)$ and $P_g(0)$. The difference observed shows that both PDFs have a non-Gaussian distribution, but the detailed shape and the scaling properties of the two PDFs are different.

In a previous section we reported that a scaling compatible with a Lévy stable process is observed for economic data and indeed a Lévy distribution reproduces quite well the central part of the distribution of the S&P 500 index variations. A similar scaling does not exist for turbulence data over a wide time interval.

The parallel analysis of the statistical properties of an economic index and the velocity of a turbulent fluid in a 3-dimensional space shows that the two processes are quantitatively different.[31]

V. DISCUSSION

In this paper we discuss some recent work done by the authors in which we analyze and model high-frequency financial data. Our conclusions are: (i) the central part of the distribution of index changes is well described by a Lévy stable distribution; (ii) extreme events (distant from the origin for more than six standard deviations approximately) are not described by a Lévy stable distribution; (iii) the scaling property of a large part of the index changes distribution and its breakdown are better detected by investigating the probability of return to the origin rather than the far tails of the measured changes distributions; (iv) a simple model, the TLF describes several (but not all) the empirical observations.

The result of point (i) is in partial agreement with the 1963 description of the stochastic nature of financial time series done by Mandelbrot[3] which states that distribution of changes of the logarithm of prices are non-Gaussian Lévy stable distributions. The disagreement between the present results and Mandelbrot's seminal results lays in the description of the far tails of the distributions. Mandelbrot describes the entire distribution as a Lévy stable distribution while we present evidence that, in addition to the presence of scaling, it is also possible to observe in high-frequency data its breakdown. Specifically, by investigating the probability of return to the origin of the stochastic process we observe breakdown of the scaling in the far tails of the index changes distributions measured with $\Delta t < 10$ minutes and we also estimate a breakdown of the scaling in time after a very long but finite interval. The observed breakdown of the scaling has two very important consequences: (a) the second moment of the index changes is finite and (b) the self-similarity of the index change distribution is only approximately true and the shape of the distribution is progressively changing from a Lévy stable shape to a Gaussian shape from short-time to long-time horizons.

We have tried to give a concrete example on how and why physicists may consider today financial systems as very interesting.

It is true that, at the moment, it may seem to be an unusual challenge for a physicist to investigate economic systems by using tools and paradigms developed to describe physical phenomena. Physical phenomena and economic systems are, of course, rather different. In physical phenomena one often may apply conservation laws and/or find equilibrium states characterized by the maximization of some (extremely relevant) extensive function as, for example, entropy. In financial systems nothing similar has been discovered yet. This makes extremely challenging the modelization of financial "complex systems."

On the other hand, physicists are today increasingly involved in research projects devoted to obtain theoretical, numerical and experimental descriptions of many-body non-equilibrium disordered (in the considered space and/or in time) systems (included non-ergodic systems). For these scholars an interdisciplinary approach to financial problems might provide a set of new problems connotated by either fundamental or applied aspects.

ACKNOWLEDGMENTS

We thank K. R. Sreenivasan for important contributions in the early stages of this work and for helpful discussions, and NSF and INFM for financial support.

REFERENCES

1. V. Pareto, *Cours d'Economie Politique* (Lausanne and Paris, 1897).
2. L. J. B. Bachelier, *Théorie de la Speculation* (Gauthier-Villars, Paris, 1900).
3. B. B. Mandelbrot, *J. Business* **36**:394 (1963); **39**:242 (1966); **40**:393 (1967).
4. E. F. Fama, *J. Business* **38**:34 (1965).
5. P. H. Cootner (ed.), *The Random Character of Stock Market Prices* (MIT Press, Cambridge, MA, 1964).
6. P. K. Clark, *Econometrica* **41**:135 (1973).
7. R. F. Engle, *Econometrica* **50**:987 (1982).
8. T. Bollerslev, R. Y. Chou, and K. F. Kroner, *J. Econometrics* **52**:5 (1992).
9. P. Bak, K. Chen, J. A. Scheinkman, and M. Woodford, *Ricerche Economiche* **47**:3 (1993).
10. R. N. Mantegna, *Physica A* **179**:232 (1991).
11. W. Li, *International Journal of Bifurcation and Chaos* **1**:583 (1991).
12. H. Takayasu, H. Miura, T. Hirabayashi, and K. Hamada, *Physica A* **184**:127 (1992).
13. J.-P. Bouchaud and D. Sornette, *J. Phys. I France* **4**:863 (1994).
14. R. N. Mantegna and H. E. Stanley, *Nature* **376**:46 (1995).
15. P. W. Anderson, K. J. Arrow, and D. Pines, *The Economy as an Evolving Complex System* (Addison-Wesley, Redwood City, 1988).
16. J. A. Scheinkman and B. J. LeBaron, *J. Business* **62**:327 (1989).

17. W. A. Brock, D. A. Hsieh, and B. J. LeBaron, *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Inference* (MIT Press, Cambridge, MA, 1991).
18. W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1971).
19. B. V. Gnedenko and A. Kolmogorov, *Limit Distribution for Sums of Independent Random Variables* (Addison-Wesley, Cambridge, MA, 1954).
20. B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1983).
21. M. F. Shlesinger, G. M. Zaslavsky, and J. Klafter, *Nature* **363**:31 (1993), and references therein.
22. P. Lévy, Théorie de l'Addition des Variables Aléatoires (Gauthier-Villars, Paris, 1937).
23. W. W. Brock and A. W. Kleidon, *Journal of Economic Dynamics and Control* **16**:451 (1990).
24. R. N. Mantegna and H. E. Stanley, in *Proceedings of the International Summer School of Physics "E. Fermi," The Physics of Complex Systems*, edited by F. Mallamace and H. E. Stanley (IOS Press, Amsterdam, 1997).
25. A. L. Tucker, *J. Bus. Econ. Stat.* **10**:73 (1992); and references therein.
26. R. N. Mantegna and H. E. Stanley, *Phys. Rev. Lett.* **73**:2946 (1994); R. N. Mantegna and H. E. Stanley, in *Lévy Flights and Related Topics in Physics*, edited by M. F. Shlesinger, G. M. Zaslavsky, and U. Frisch (Springer, Berlin, 1995), p. 300.
27. M. F. Shlesinger, *Phys. Rev. Lett.* **74**:4959 (1995).
28. I. Koponen, *Phys. Rev.* **E52**:1197 (1995).
29. R. N. Mantegna, *Phys. Rev.* **E49**:4677 (1994).
30. G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance* (Chapman and Hall, NY, 1994).
31. R. N. Mantegna and H. E. Stanley, *Nature* **383**:587 (1996); *Physica* **239**:225 (1997).
32. G. W. Schwert, *The Journal of Finance* **XLIV**:1115 (1989).
33. U. Frisch, *Turbulence* (Cambridge Univ. Press, Cambridge, UK, 1995).
34. S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge, *Nature* **381**:767 (1996).
35. A. Arneodo *et al.*, COND-MAT 9607120.
36. P. Kailasnath, K. R. Sreenivasan, and G. Stolovitzky, *Phys. Rev. Lett.* **68**:2766 (1992).