67

12. Investigations of Financial Markets Using Statistical Physics Methods

Rosario N. Mantegna and H. Eugene Stanley

We begin with a brief historical note concerning the growing interest of statistical physicists in the analysis and modeling of financial markets. We then briefly discuss the key concepts of arbitrage and efficient markets. We relate these concepts to apparently 'universal' aspects observed in the empirical analysis of stock price dynamics in financial markets. In particular, we consider (i) the empirical behavior of the probability density function for the return of an economic time series to where it started and (ii) the content of economic information in a financial time series.

12.1 Introduction

The quantitative modeling of financial markets started in 1900 with the pioneering work of the French mathematician Bachelier [12.1]. Since the 1950s, the analysis and modeling of financial markets have become an important research area of economics and financial mathematics [12.2]. The researches pursued have been very successful, and nowadays a robust theoretical framework char-

NEXT PAGE

Fig. 12.0. Color representation quantifying the stability in time of the eigenvectors of the correlation matrix that deviate from random-matrix bounds. Two partially overlapping time periods A and B, of four months each, were analyzed, January 1994 – April 1994 and March 1994 – June 1994. Each of the 225 squares has a rainbow color proportional to the scalar product ('overlap') of the largest 15 eigenvectors of the correlation matrix in period A with those of the same 15 eigenvectors from period B. Perfect stability in time would imply that this pixel representation of the overlaps has ones (the red end of the rainbow spectrum) in the diagonal and zeros (violet) in the off-diagonal. The eigenvectors are shown in inverse rank order (from smallest to largest), and we note that the pixels near the upper right corner have colors near the red end of the spectrum, corresponding to the fact that the largest 6–8 eigenvectors are relatively stable; in particular, the largest 3–4 eigenvectors are stable for very long periods of time. The remainder of the pixels are distributed toward the violet end of the spectrum, corresponding to the fact that the overlaps are not statistically significant, and corroborating the finding that their corresponding eigenvalues are random. This figure is kindly contributed by P. Gopikrishnan and V. Plerou

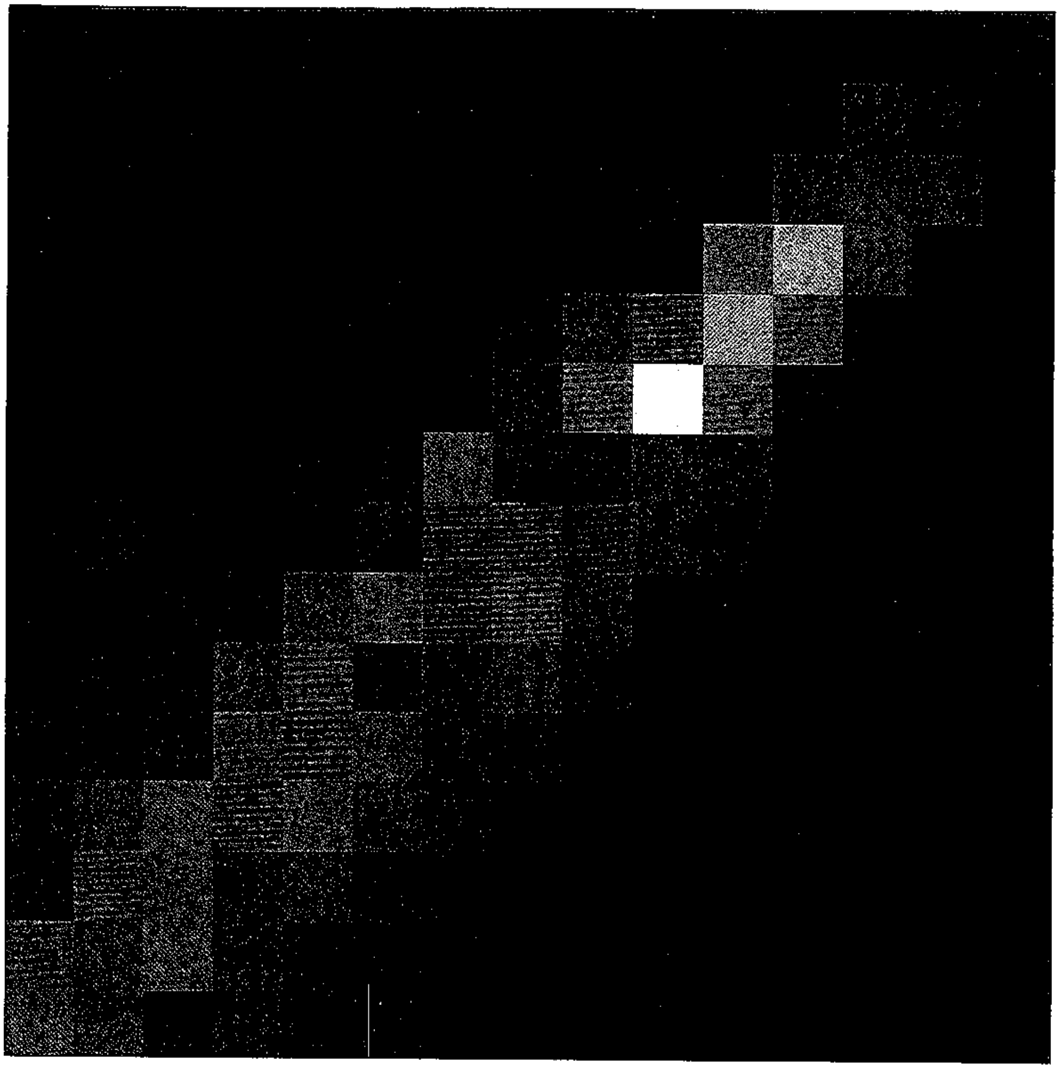


FIG. 12.0

acterizes these disciplines [12.3–12.6]. In parallel to these studies, starting from the 1990s a group of physicists became interested in the analysis and modeling of financial markets by using tools and paradigms of their own discipline (for an overview, consider, for example, [12.7–12.10]). The interest of physicists in such systems is directly related to the fact that, during the years, predictability has assumed a meaning in physics, which is quite different from the one originally associated with the predictability of, for example, a Newtonian linear system. The degree of predictability of physics systems is nowadays known to be essentially limited in nonlinear and complex systems. This makes the physical prediction less strong, but on the other hand the area of research covered by physical investigations and of its application may increase [12.11].

In addition to the above observations, there are a series of reasons explaining why this discipline emerged during the last decade. We will try to discuss some of them hereafter.

Since the 1970s, a series of significant changes has taken place in the world of finance. One key year was 1973, when currencies began to be traded in financial markets and their values determined by the foreign exchange market, a financial market active 24 hr a day all over the world. During that same year, Black and Scholes [12.12] published the first paper that presented a rational option-pricing formula.

Since that time, the volume of foreign exchange trading has been growing at an impressive rate. The transaction volume in 1995 was 80 times what it was in 1973. An even more impressive growth has taken place in the field of derivative products. The notional amount of financial derivative market contracts issued in 1999 was 81 trillion US dollars. Contracts were negotiated in the over-the-counter market (i.e. directly between firms or financial institutions), and in specialized exchanges that deal only in derivative contracts. Today, financial markets facilitate the trading of huge amounts of money, assets, and goods in a competitive global environment.

A second revolution began in the 1980s when electronic trading was adapted to the foreign exchange market. The electronic storing of data relating to financial contracts - or to bid and ask quotes issued by traders - was put in place at about the same time that electronic trading became widespread. One result is that today a huge amount of electronically stored financial data is readily available. These data are characterized by the property of being high-frequency data - the average time delay between two records can be as short as a few seconds. Between the available databases it may be worth mentioning the Olsen and Associates database comprising all the bid and ask quotes of the foreign exchange market collected from information vendors, including Reuters, Knight-Ridder, and Telerate since 1986 and the Trade and Quote (TAQ) database, which comprises all the trades and quotes related to all the securities listed in the New York Stock Exchange, Nasdaq National Market System, and SmallCap issues. The TAQ database is available monthly on CD-

Rom from the New York Stock Exchange. The enormous expansion of financial markets requires strong investments in money and human resources to achieve reliable quantification and minimization of risk for the financial institutions involved.

12.2 Econophysics

The research approach of physicists to financial modeling aims to be complementary to the ones of financial mathematicians and economists. The main goals are (i) to contribute to a better understanding and modeling of financial markets and (ii) to promote the use of physical concepts and expertise in the multidisciplinary approach to risk management.

This research area is often addressed as *econophysics*. The word econophysics describes the present attempts of a number of physicists to model financial and economic systems using paradigms and tools borrowed from theoretical and statistical physics.

Financial markets exhibit several of the properties that characterize complex systems. They are open systems in which many subunits interact nonlinearly in the presence of feedback. In financial markets, the governing rules are rather stable and the time evolution of the system is continuously monitored. It is now possible to develop models and to test their accuracy and predictive power using available data, since large databases exist even for high-frequency data.

A research community has begun to emerge in econophysics starting from the 1990s. New interdisciplinary journals have been published, conferences have been organized, and a set of potentially tractable scientific problems has been provisionally identified. The research activity of this group of physicists is complementary to the most traditional approaches of finance and mathematical finance. One characteristic difference is the emphasis that physicists put on the empirical analysis of economic data. Another is the background of theory and method in the field of statistical physics developed over the past 30 yr that physicists bring to the subject. The concepts of scaling, universality, disordered frustrated systems, and self-organized systems might be helpful in the analysis and modeling of financial and economic systems. One argument that is sometimes raised at this point is that an empirical analysis performed on financial or economic data is not equivalent to the usual experimental investigation that takes place in physical sciences. In other words, it is impossible to perform large-scale experiments in economics and finance that could falsify any given theory.

We note that this limitation is not specific to economic and financial systems, but also affects such well developed areas of physics as astrophysics, atmospheric physics, and geophysics. Hence, in analogy to activity in these more established areas, we find that we are able to test and falsify any theories associated with the currently available sets of financial and economic data provided in the form of recorded files of financial and economic activity.

Among the important areas of physics research dealing with financial and economic systems, one concerns the complete statistical characterization of the stochastic process of price changes of a financial asset. Several studies have been performed that focus on different aspects of the analyzed stochastic process, e.g. the shape of the distribution of price changes [12.9, 12.13–12.17], the temporal memory [12.18–12.21], and the higher-order statistical properties [12.22–12.24]. This is still an active area, and attempts are ongoing to develop the most satisfactory stochastic model describing all the features encountered in empirical analyses. One important accomplishment in this area is an almost complete consensus concerning the finiteness of the second moment of price changes. This has been a longstanding problem in finance, and its resolution has come about because of the renewed interest in the empirical study of financial systems.

A second area concerns the development of a theoretical model that is able to encompass all the essential features of real financial markets. Several models have been proposed [12.25–12.39], and some of the main properties of the stochastic dynamics of stock price are reproduced by these models as, for example, the leptokurtic non-Gaussian shape of the distribution of price differences. Parallel attempts in the modeling of financial markets by taking into account some results observed in the empirical analyses have been also developed by economists [12.40–12.42, 12.87].

One of the more active areas in finance is the pricing of derivative instruments. In the simplest case, an asset is described by a stochastic process and a derivative security (or contingent claim) is evaluated on the basis of the type of security and the value and statistical properties of the underlying asset. This problem presents at least two different aspects: (i) 'fundamental' aspects, which are related to the nature of the random process of the asset, and (ii) 'applied' or 'technical' aspects, which are related to the solution of the option-pricing problem under the assumption that the underlying asset performs the proposed random process.

In this area the investigations which are considering the problem of the rational pricing of a derivative product when some of the canonical assumptions of the Black and Scholes model are relaxed [12.9, 12.43, 12.44]. Other autors focus on aspects of portfolio selection and its dynamical optimization [12.45–12.49]. A further area of research considers analogies and differences between price dynamics in a financial market and such processes as turbulence [12.15, 12.19, 12.50] and the dynamics of ecological systems [12.16, 12.51].

Another common theme encountered in econophysics concerns the time correlation of financial series. The detection of the presence of a higher-order correlation in price changes has motivated a reconsideration of some beliefs of what is termed 'technical analysis' [12.52].

In addition to the studies that analyze and model financial systems, there are studies of the income distribution of firms and studies of the statistical properties of their growth rates [12.53–12.56]. The statistical properties of the economic performances of complex organizations such as universities or entire countries have also been investigated [12.57].

This brief presentation of some of the current efforts in this emerging discipline has only illustrative purposes and cannot be exhaustive. For a more complete overview, consider, for example, the proceedings of conferences dedicated to these topics [12.7, 12.8].

From the above overview we see that econophysics started as an emerging discipline during the 1990s. However before the starting up of the discipline a series of physicists and mathematicians have investigated on an individual basis financial and economic problems. Some of these pioneering approaches are discussed in the next section.

12.3 An Historical Note

The interest of the physics community in financial and economic systems has roots that date back to 1942, when a paper from Majorana on the essential analogy between statistical laws in physics and in the social sciences was published [12.58]. In his contribution, he wrote that "......It is important that the principles of quantum mechanics have led to recognize (...) the statistical character of the fundamental laws of elementary processes. This conclusion makes essential the analogy between physics and social sciences, between which there is an identity of values and method". This unorthodox point of view was considered of marginal interest until recently. Indeed, prior to the 1990s, very few professional physicists did any research associated with social or economic systems. The exceptions included Kadanoff [12.59], Montroll and Badger [12.60], and a group of physical scientists at the Santa Fe Institute [12.61].

In this chapter we briefly discuss the application to financial markets of such concepts as power-law distributions, correlations, scaling, unpredictable time series, and random processes. During the past 30 yr, physicists have achieved important results in the field of phase transitions, statistical mechanics, nonlinear dynamics, and disordered 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 sciences to learn that the first use of a power-law distribution - and the first mathematical formalization of a random walk - took place in the social sciences. In 1897 the Italian social economist Pareto investigated the statistical character of the wealth of individuals in a stable economy by modeling them using the distribution

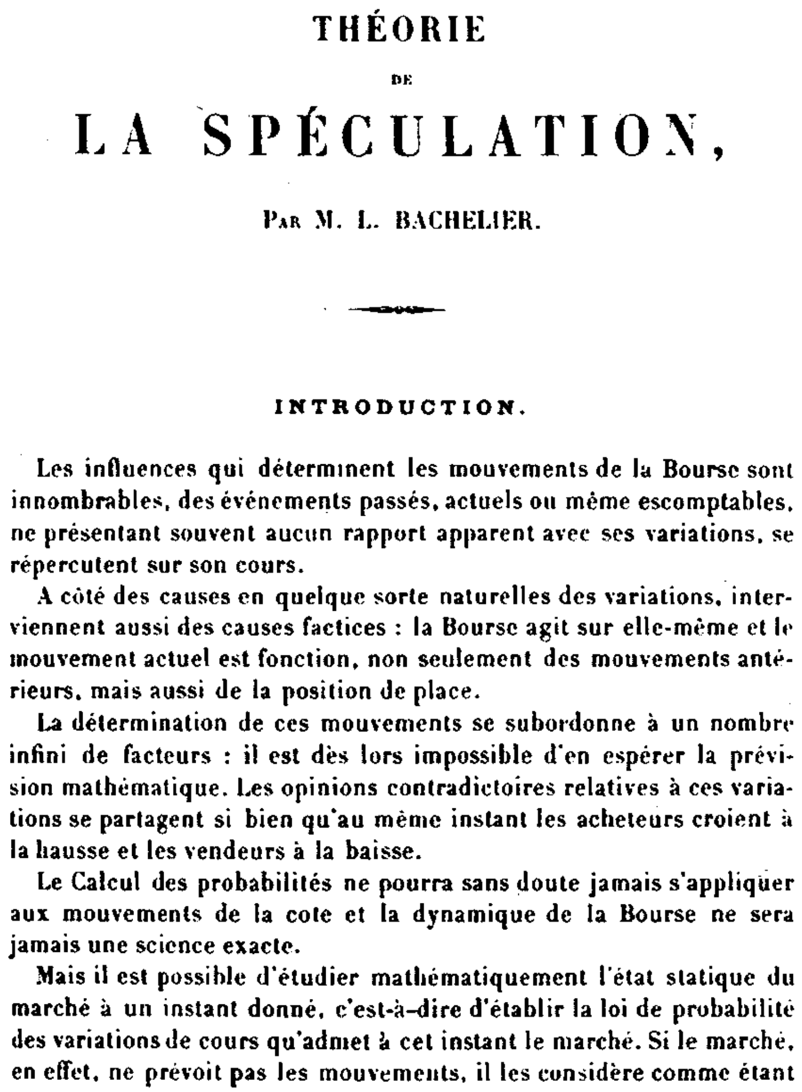
$$y \sim x^{-\nu}, \tag{12.1}$$

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 [12.62]. 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 counterintuitive, because it 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 Lévy [12.63] and thanks to the application of power-law distributions to several problems pursued by Mandelbrot [12.64]; and (ii) in the study of phase transitions, which introduced the concepts of scaling for thermodynamic functions and correlation functions [12.65].

Another concept ubiquitous in the natural sciences is the random walk. The first theoretical description of a random walk in the natural sciences was performed in 1905 by Einstein [12.66] in his famous paper dealing with the determination of the Avogadro number. In subsequent years, the mathematics of the random walk was made more rigorous by Wiener [12.67], and now the random walk concept has spread across almost all research areas in the natural sciences.

The first formalization of a random walk was not in a publication by Einstein, but in the doctoral thesis by Bachelier [12.1]. Bachelier presented his thesis to the faculty of sciences at the Academy of Paris on 29 March 1900, for the degree of *Docteur en Sciences Mathématiques.* The first page of his thesis is shown in Fig. 12.1. His advisor was Poincaré, one of the greatest mathematicians of his time. The thesis, entitled *Théorie de la Spéculation*, is surprising in several respects. It deals with the pricing of options in speculative markets, an activity that today is extremely important in financial markets where derivative securities - those whose value depends on the values of other more basic underlying variables - are regularly traded on many different exchanges. To complete this task, Bachelier determined the probability of price changes by writing down what is now called the Chapman-Kolmogorov equation and recognizing that what is now called a Wiener process satisfies the diffusion equation (this point was re-discovered by Einstein in his 1905 paper on Brownian motion). Retrospectively analyzed, Bachelier's thesis lacks rigor in some of its mathematical and economic points. Specifically, the determination of a Gaussian distribution for the price changes was - mathematically speaking - not sufficiently motivated. On the economic side, Bachelier investigated price changes, whereas economists are mainly dealing with changes in



THÉORIE

DE

LA SPÉCULATION,

PAR M. L. BACHELIER.

INTRODUCTION.

Les influences qui déterminent les mouvements de la Bourse sont innombrables, des événements passés, actuels ou même escomptables, ne présentant souvent aucun rapport apparent avec ses variations, se répercutent sur son cours.

A côté des causes en quelque sorte naturelles des variations, interviennent aussi des causes factices : la Bourse agit sur elle-même et le mouvement actuel est fonction, non seulement des mouvements antérieurs, mais aussi de la position de place.

La détermination de ces mouvements se subordonne à un nombre infini de facteurs : il est dès lors impossible d'en espérer la prévision mathématique. Les opinions contradictoires relatives à ces variations se partagent si bien qu'au même instant les acheteurs croient à la hausse et les vendeurs à la baisse.

Le Calcul des probabilités ne pourra sans doute jamais s'appliquer aux mouvements de la cote et la dynamique de la Bourse ne sera jamais une science exacte.

Mais il est possible d'étudier mathématiquement l'état statique du marché à un instant donné, c'est-à-dire d'établir la loi de probabilité des variations de cours qu'admet à cet instant le marché. Si le marché, en effet, ne prévoit pas les mouvements, il les considère comme étant

Fig. 12.1. First page of the PhD thesis of Luis Bachelier. This work was written in 1900, almost 50 yr before the concept of random walk was currently used to model asset price dynamics in a financial market and 73 yr before the first publication of a rational option price procedure

the logarithm of price. However, these limitations do not diminish the value of Bachelier's pioneering work.

To put Bachelier's work into perspective, the Black and Scholes option-pricing model - considered the milestone in option-pricing theory - was published in 1973, almost three-quarters of a century after the publication of his thesis. Moreover, theorists and practitioners are aware that the Black and Scholes model needs correction in its application, meaning that the problem of which stochastic process describes the changes in the logarithm of prices in a financial market is still an open one.

The problem of the distribution of price changes has been considered by several authors since the 1950s, which was the period when mathematicians began to show interest in the modeling of stock market prices. Bachelier's original proposal of Gaussian distributed price changes was soon replaced by a model in which stock prices are log-normal distributed, i.e. stock prices are performing a geometric Brownian motion. In a geometric Brownian motion, the differences of the logarithms of prices are Gaussian distributed. This model is known to provide only a first approximation of what is observed in real data. For this reason, a number of alternative models have been proposed with the aim of explaining

(i) the empirical evidence that the tails of measured distributions are fatter than expected for a geometric Brownian motion; and
(ii) the time fluctuations of the second moment of price changes.

Among the alternative models proposed, 'the most revolutionary development in the theory of speculative prices since Bachelier's initial work' [12.2], is Mandelbrot's hypothesis that price changes follow a Lévy stable distribution [12.68]. Lévy stable processes are stochastic processes obeying a generalized central limit theorem. By obeying a generalized form of the central limit theorem, they have a number of interesting properties. They are stable (as are the more common Gaussian processes) - i.e. the sum of two independent stochastic processes x_1 and x_2 characterized by the same Lévy distribution of index α is itself a stochastic process characterized by a Lévy distribution of the same index. The shape of the distribution is maintained (is stable) by summing up independent and identically distributed Lévy stable random variables.

Lévy stable processes define a basin of attraction in the functional space of probability density functions. The sum of independent and identically distributed stochastic processes $S_n \equiv \sum_{i=1}^{n} x_i$ characterized by a probability density function with power-law tails,

$$P(x) \sim x^{-(1+\alpha)}, \tag{12.2}$$

will converge, in probability, to a Lévy stable stochastic process of index α when n tends to infinity [12.69].

This property tells us that the distribution of a Lévy stable process is a power-law distribution for large values of the stochastic variable x. The fact that power-law distributions lack a typical scale is reflected in Lévy stable processes by the property that the variance of Lévy stable processes is infinite for $\alpha < 2$. Stochastic processes with infinite variance, although well defined mathematically, are extremely difficult to use and, moreover, raise fundamental questions when applied to real systems. For example, in physical systems the second moment is often related to the system temperature, so infinite variances imply an infinite (or undefined) temperature. In financial systems, an infinite variance would complicate the important task of risk estimation.

12.4 Key Concepts

In the model of financial markets a pair of concepts are crucial to understand the current theories modeling them. In this section we will discuss two relevant concepts namely the concept of absence of arbitrage and the concept of efficient market. Financial markets are systems in which a large number of traders interact with one another and react to external information in order to determine the best price for a given asset. The goods might be as different as animals, ore, equities, currencies, or bonds - or derivative products issued on those underlying financial goods. Some markets are localized in specific cities (e.g. New York, Tokyo, and London) while others (such as the foreign exchange market) are delocalized and accessible all over the world.

When one inspects a time series of the time evolution of the price, volume, and number of transactions of a financial product, one recognizes that the time evolution is unpredictable. At first sight, one might sense a curious paradox. An important time series, such as the price of a financial good, is essentially indistinguishable from a stochastic process. There are deep reasons for this kind of behavior, and in this chapter we will examine some of these.

12.4.1 Arbitrage

A key concept for the understanding of markets is the concept of arbitrage - the simultaneous purchase and sale of the same or equivalent security in order to profit from price discrepancies.

The presence of traders looking for arbitrage conditions contributes to a market's ability to evolve the most rational price for a good. To see this, suppose that one has discovered an arbitrage opportunity. One will exploit it and, if one succeeds in making a profit, one will repeat the same action. This action will increase the demand of a given good at a given place or time and will simultaneously increase the supply of the same good at another time or place. The modification of the demand and supply levels forces the price of the considered good to attain a more rational value in a given place or time.

To summarize: (i) new arbitrage opportunities continually appear and are discovered in the markets but (ii) as soon as an arbitrage opportunity begins to be exploited, the system moves in a direction that gradually eliminates the arbitrage opportunity.

12.4.2 Efficient Market Hypothesis

Markets are complex systems that incorporate information about a given asset in the time series of its price. The most accepted paradigm among scholars in finance is that the market is highly efficient in the determination of the most rational price of the traded asset. The efficient market hypothesis (EMH) was originally formulated in the 1960s [12.70]. A market is said to be efficient if all the available information is instantly processed when it reaches the market and it is immediately reflected in a new value of prices of the assets traded.

The theoretical motivation for the efficient market hypothesis has its roots in the pioneering work of Bachelier [12.1], who proposed that the price of assets in a speculative market be described as a stochastic process. This work remained almost unknown until the 1950s, when empirical results [12.2] about the serial correlation of the rate of return showed that correlations on a short time scale are negligible and that the approximate behavior of return time series is indeed similar to uncorrelated random walks.

The EMH was formulated explicitly in 1965 by Samuelson [12.71], who showed mathematically that properly anticipated prices fluctuate randomly. Using the hypothesis of rational behavior and market efficiency, he was able to demonstrate how Y_{t+1}, the expected value of the price of a given asset at time $t+1$, is related to the previous values of prices $Y_0, Y_1, \dots, Y_t$ through the relation

$$E\{Y_{t+1}|Y_0, Y_1, \dots, Y_t\} = Y_t. \tag{12.3}$$

Stochastic processes obeying the conditional probability given in (12.3) are called martingales [12.72]. The notion of a martingale is, intuitively, a probabilistic model of a 'fair' game. In gambler's terms, the game is fair when gains and losses cancel, and the gambler's expected future wealth coincides with the gambler's present assets. The fair game conclusion about the price changes observed in a financial market is equivalent to the statement that there is no way of making a profit on an asset by simply using the recorded history of its price fluctuations. The conclusion of this 'weak form' of the EMH is then that price changes are unpredictable from the historical time series of those changes.

Since the 1960s, a great number of empirical investigations have been devoted to testing the efficient market hypothesis [12.73]. In the great majority of the empirical studies, the time correlation between price changes has been found to be negligibly small, supporting the efficient market hypothesis. However, it was shown in the 1980s that by using the information present in additional time series such as earnings/price ratios, dividend yields, and term-structure variables, it is possible to make predictions of the rate of return of a given asset on a long time scale, much longer than a month. Thus empirical observations have challenged the stricter form of the efficient market hypothesis.

Thus empirical observations and theoretical considerations show that price changes are difficult if not impossible to predict if one starts from the time series of price changes. In its strict form, an efficient market is an idealized system. In actual markets, residual inefficiencies are always present. Searching out and exploiting arbitrage opportunities is one way of eliminating market inefficiencies.

12.5 Idealized Systems in Physics and Finance

The efficient market is an idealized system. Real markets are only approximately efficient. This fact will probably not sound too unfamiliar to physicists because they are well acquainted with the study of idealized systems. Indeed, the use of idealized systems in scientific investigation has been instrumental in the development of physics as a discipline. Where would physics be without idealizations such as frictionless motion, reversible transformations in thermodynamics, and infinite systems in the critical state? Physicists use these abstractions in order to develop theories and to design experiments. At the same time, physicists always remember that idealized systems only approximate real systems, and that the behavior of real systems will always deviate from that of idealized systems. A similar approach can be taken in the study of financial systems. We can assume realistic 'ideal' conditions, e.g. the existence of a perfectly efficient market, and within this ideal framework develop theories and perform empirical tests. The validity of the results will depend on the validity of the assumptions made.

The concept of the efficient market is useful in any attempt to model financial markets. After accepting this paradigm, an important step is to fully characterize the statistical properties of the random processes observed in financial markets.

12.6 Empirical Analysis

Econophysics is an interdisciplinary research area, with a growing number of practitioners. We shall briefly describe the spirit and substance of some recent work that focuses on universal aspects observed in the empirical analysis of different (in location and time period) financial markets.

12.6.1 Statistical Properties of Price Dynamics

The knowledge of the statistical properties of price dynamics in financial markets is fundamental. It is necessary for any theoretical modeling aiming to obtain a rational price for a derivative product issue on it [12.74] and it is the starting point of any valuation of the risk associated with a financial position [12.75]. Moreover, it is needed in any effort aiming to model the system. In spite of this importance, the modeling of such variable is not yet conclusive.

Several models exist which are showing partial successes and unavoidable limitations. In this research, the approach of physicists maintains the specificity of their discipline namely to develop and modify models by taking into account results of empirical analysis.

Several models have been proposed and we will not review them here. Here we wish only to focus on the aspects which are *universally* observed in various stock price and index price dynamics.

12.6.2 Short- and Long-Range Correlations

In any financial market - either well established and highly active as the New York stock exchange, *emerging* as the Budapest stock exchange, or *regional* as the Milan stock exchange - the autocorrelation function of returns is a monotonic decreasing function with a very short correlation time. High frequency data analyses have shown that correlation times can be as short as a few minutes in highly traded stocks or indices [12.19, 12.76]. A fast decaying autocorrelation function is also observed in the empirical analysis of data recorded transaction by transaction. By using as time index the number of transactions occurred from a selected origin, a time memory as short as a few transactions has been detected in the dynamics of most traded stocks of the Budapest 'emerging' financial market [12.77].

The short-range memory between returns is directly related to the necessity of absence of continuous arbitrage opportunities in efficient financial markets. In other words, if correlation were present between returns this would allow devising trading strategies that would provide a net gain continuously and without risk. The continuous search for and exploitation of arbitrage opportunities from traders focused on this kind of activity drastically diminish the redundancy in the time series of price changes. Another mechanism reducing the redundancy of stock price time series is related to the presence of so-called 'noise traders'. With their action, noise traders add into the time series of stock price information, which is unrelated to the economic information decreasing the degree of redundancy of the price changes time series.

It is worth pointing out that not all the economic information present in stock price time series disappears due to these mechanisms. Indeed the redundancy that needs to be eliminated concerns only price change and not any nonlinear functions of it [12.78].

The absence of time correlation between returns does not mean that returns are identically distributed over time. In fact different authors have observed that nonlinear functions of return such as the absolute value or the square are correlated over a time scale much longer than a trading day. Moreover the functional form of this correlation seems to be power-law up to at least 20 trading days approximately [12.20, 12.21, 12.24, 12.76, 12.79–12.81].

A final observation concerns the degree of stationary behavior of the stock price dynamics. Empirical analysis shows that returns are not strictly sense stationary stochastic processes. Indeed the volatility (standard deviation of returns) is itself a stochastic process. Although a general proof is still lacking, empirical analyses performed on financial data of different financial markets suggest that the stochastic process is locally nonstationary but asymptotically stationary. By asymptotically stationary we mean that the probability density function (PDF) of returns measured over a wide time interval exists and it is uniquely defined. A paradigmatic example of simple stochastic processes which are locally nonstationary but asymptotically stationary is provided by ARCH [12.82] and GARCH [12.83] processes.

12.6.3 The Distribution of Returns

The PDF of returns shows some 'universal' aspects. By 'universal' aspects we mean that they are observed in different financial markets at different periods of time provided that a sufficiently long time period is used in the empirical analysis. The first of these 'universal' or stylized facts is the leptokurtic nature of the PDF. A leptokurtic PDF characterizes a stochastic process having small changes and very large changes more frequent than in the case of Gaussian distributed changes. Leptokurtic PDFs have been observed in stocks and indices time series by analyzing both high-frequency and daily data. An example is provided in Fig. 12.2. Thanks to the recent availability of transaction-by-transaction data, empirical analyses on a transaction time scale have also been performed. One of these studies performed by analyzing stock price in the Budapest stock exchange show that return PDF are leptokurtic down to a 'transaction' time scale [12.77]. See a direct example in Fig. 12.3.

The origin of the observed leptokurtosis is still debated. There are several models trying to explain it. Just to cite (rather arbitrarily) a few of them: (i) a model of Lévy stable stochastic process [12.68]; (ii) a model assuming that the non-Gaussian behavior occurs as a result of the uneven activity during market hours [12.84]; (iii) a model where a geometric diffusive behavior is superimposed to Poissonian jumps [12.85]; (iv) a quasi-stable stochastic process with finite variance [12.86]; and (v) a stochastic process with rare events described by a power-law exponent not falling into the Lévy regime [12.17, 12.41, 12.87]. The above processes are characterized by finite or infinite moments. In the attempt to find the stochastic process that best describes stock price dynamics, it is then important to try to preliminary conclude about the finiteness or infiniteness of the second moment.

The above answer is not simply obtained [12.88] and careful empirical analyses must be performed to reach a reliable conclusion. It is our opinion that an impressive amount of empirical evidence has been recently found supporting the conclusion that the second moment of the return PDF is fi-

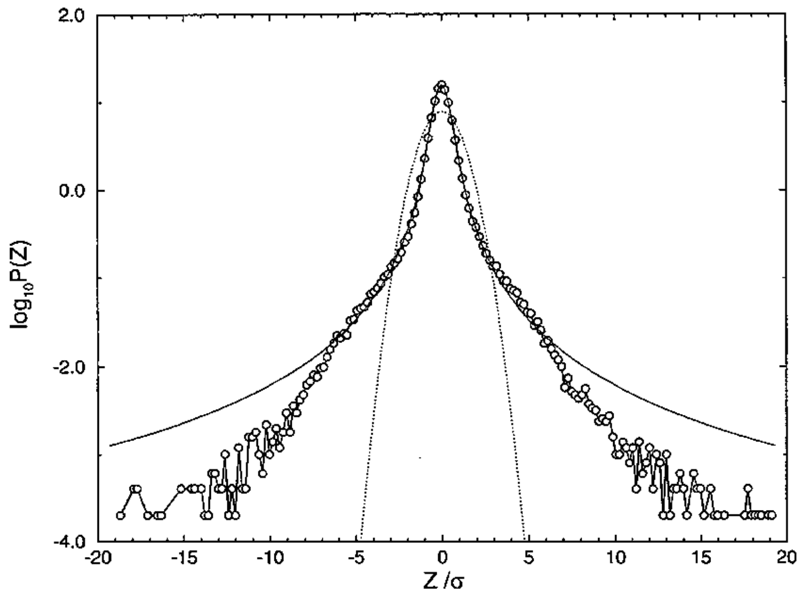


Fig. 12.2. Logarithm of the probability density function of the S&P 500 index high-frequency changes computed at a $\Delta t = 1$ min time horizon. The probability density function is compared with the symmetrical Lévy stable distribution of index $\alpha = 1.40$ and scale factor $\gamma = 0.00375$ (solid line). The dotted line is the Gaussian distribution with standard deviation σ equal to the experimental value 0.0508. The variations of price are normalized to this value. Deviations from the Lévy stable profile are observed for $|z|/\sigma \gtrsim 6$. From [12.14]

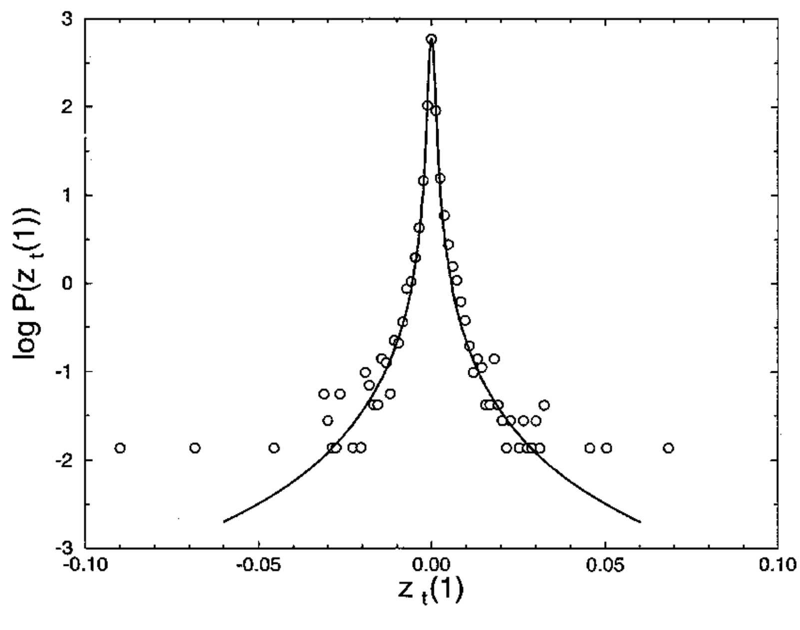


Fig. 12.3. Semilogarithmic plot of the probability density function of tick by tick log price changes measured for the MOL Company (the time period is the second quarter of 1998). The solid line is a symmetrical Lévy distribution of index $\alpha = 1.60$ and scale factor $\gamma = 0.000005$. From [12.77]

nite [12.14, 12.17, 12.41, 12.87, 12.89, 12.90]. This conclusion has a deep consequence on the stability of the return PDF. The finiteness of the second moment and the independence of successive returns imply that the central limit theorem asymptotically applies. Hence the form expected for the return PDF must be Gaussian for very long time horizons. We then have two regions - at short time horizons we observe leptokurtic distributions whereas at long time horizons we expect a Gaussian distribution. A complete characterization of the stochastic process needs an investigation performed at different time horizons. During this kind of analysis, non-Gaussian scaling and its breakdown has been detected [12.14, 12.19].

12.7 Collective Dynamics

In the previous sections we saw that 'universal' facts suggest that the stock price change dynamics in financial markets is well described by an unpredictable time series. However, this does not imply that the stochastic dynamics of stock price time series is a random walk with independent identically distributed increments. Indeed the stochastic process is much more complex than a customary random walk.

One key question in the analysis and modeling of a financial market concerns the independence of the price time series of different stocks traded simultaneously in the same market. The presence of cross-correlations between pairs of stocks has been known for a long time and it is one of the basic assumptions of the theory of the selection of the most efficient portfolio of stocks [12.91]. Recently, physicists have also started to investigate empirically and theoretically the presence of such cross-correlations.

It has been found that a meaningful economic taxonomy may be obtained by starting from the information stored in the time series of stock price only. A simple example is shown in Fig. 12.4. This has been achieved by assuming that a metric distance can be defined between the synchronous time evolution of a set of stocks traded in a financial market and under the essential *Ansatz* that the subdominant ultrametric associated with the selected metric distance is controlled by the most important economic information stored in the time evolution dynamics [12.92].

Another approach to detect collective movement of stock price fluctuations involves the comparison of the statistics of the measured equal-time cross-correlation matrix C against a 'null hypothesis' of a 'random' cross-correlation matrix, constructed from mutually uncorrelated time series [12.93, 12.94]. The comparison is performed in the diagonal basis, and it is found that $\approx 98\%$ of the eigenvalues of C are consistent [12.93, 12.94] with that of a random cross-correlation matrix (see Fig. 12.5). There are also deviations [12.93, 12.94] for $\approx 2\%$ of the eigenvalues at both edges of the eigenvalue spectrum, which

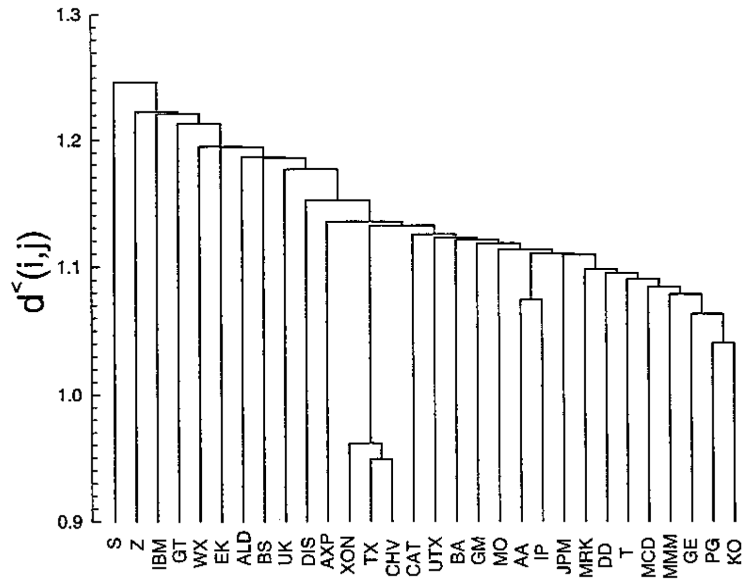


Fig. 12.4. Hierarchical tree associated to the stock portfolio of 30 stocks used to compute the Dow Jones Industrial Average index in 1998. The 30 stocks are labeled by their tick symbols. The hierarchical tree is obtained with the correlation based clustering procedure of [12.92]. In the hierarchical tree, several groups of stocks are detected. They are homogeneous with respect to the economic activities of the companies: (i) oil companies (Exxon (XON), Texaco (TX), and Chevron (CHV)); (ii) raw material companies (Alcoa (AA) and International paper (IP)), and (iii) companies working in the sectors of consumer nondurable products (Procter and Gamble (PG)) and food and drinks (Coca Cola (KO)). The distance between each stock and the others is the ultrametric distance $d^<(i,j)$ computed starting from the correlation coefficient matrix. From [12.92]

are found to correspond mainly to conventionally identified sectors of business activity [12.96].

The observation of the presence of a certain degree of statistical synchrony in the stock price dynamics suggests the following conclusion. Consideration of the time evolution of only a single stock price could be insufficient to reach a complete modeling of all essential aspects of a financial market.

12.8 Conclusion

This chapter briefly discusses the goals and scopes of econophysics, the motivations and precursors of physicists involved in the analysis and modeling of financial markets, and some of the stylized 'universal' facts that are observed in financial markets and are considered robust by several researchers working in the field. Starting from these results one can devise studies trying to enrich and expand this knowledge to provide theoreticians and computer scientists with the empirical facts that need to be explained by their models progressively proposed. The ultimate goal is to contribute to the search for the best model describing a financial market, one of the most intriguing *complex systems*.

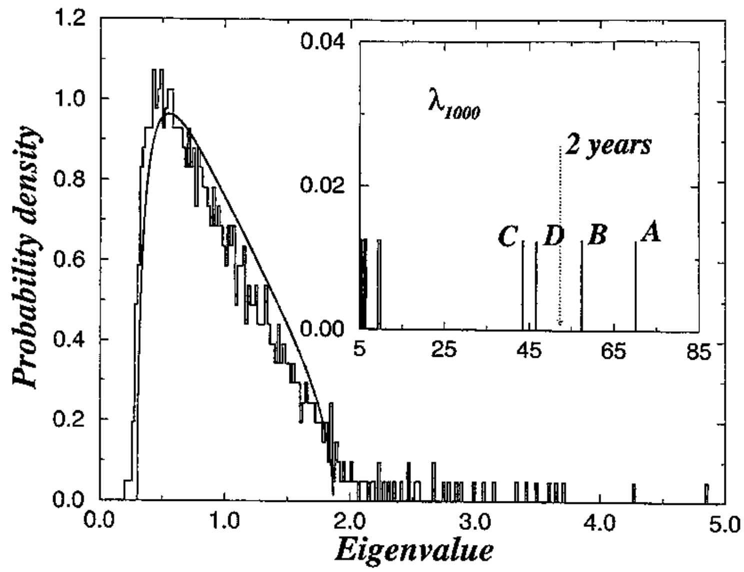


Fig. 12.5. The probability density of the eigenvalues of the equal-time cross-correlation matrix C constructed from price fluctuations of 1 000 largest stocks in the TAQ database for the 2 yr period 1994–1995. Recent analytical results [12.95] for cross-correlation matrices constructed from mutually uncorrelated time series predict a distribution of eigenvalues within a finite range depending on the ratio R of the length of the time series to the dimension of the matrix (solid curve). In our case $R = 6.448$ corresponding to eigenvalues distributed in the interval $0.37 \leq \lambda_k \leq 1.94$ [12.93–12.95]. However, the largest eigenvalue for the 2 yr period (inset) is approximately 30 times larger than the maximum eigenvalue predicted for uncorrelated time series. The inset also shows the largest eigenvalue for the cross-correlation matrix for four half-yr periods - denoted A, B, C, D. The arrow in the inset corresponds to the largest eigenvalue for the entire 2 yr period, $\lambda_{1\,000} \approx 50$. The distribution of eigenvector components for the large eigenvalues, well outside the bulk, shows significant deviations from the Gaussian prediction of RMT, which suggests 'collective' behavior or correlations between different companies. The largest eigenvalue corresponds to the correlations within the entire market [12.93, 12.94]

Acknowledgements. We thank our many collaborators for joining us in work reported herein. These include L.A.N. Amaral, G. Bonanno, P. Cizeau, P. Gopikrishnan, F. Lillo, Y. Liu, M. Meyer, Z. Palagyi, V. Plerou, and B. Rosenow. Finally, we thank INFM, MURST, and NSF for financial support.

References

12.1 L. Bachelier, Ann. Sci. l'Ecole Norm. Supér. **III-17**, 21 (1900).
12.2 P.H. Cootner (ed.), *The Random Character of Stock Market Prices* (MIT, Cambridge, 1964).
12.3 J.E. Ingersoll Jr., *Theory of Financial Decision Making* (Rowman and Littlefield, Savage, 1987).
12.4 R.C. Merton, *Continuous-Time Finance* (Blackwell, Cambridge, 1990).
12.5 I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus, 2nd edn.* (Springer, Berlin, 1991).
12.6 J.Y. Campbell, A.W. Lo, and A.C. MacKinlay, *The Econometrics of Financial Markets,* (Princeton University Press, Princeton, 1997).

12.7 R.N. Mantegna (ed.), *Proceedings of the International Workshop on Econophysics and Statistical Finance*, Physica A [special issue] **269** (1999).
12.8 J.-P. Bouchaud, K. Lauritsen, and P. Alstrom (eds.), *Proceedings of the International Conference on Applications in Finacial Analysis*, Int. J. Theor. Appl. Finance [special issue] **3**(3), 309 (2000).
12.9 J.-P. Bouchaud and M. Potters, *Theory of Financial Risk* (Cambridge University Press, Cambridge, 2000).
12.10 R.N. Mantegna and H.E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge, 2000).
12.11 G. Parisi, Physica A **263**, 557 (1999).
12.12 F. Black and M. Scholes, J. Polit. Econ. **81**, 637 (1973).
12.13 R.N. Mantegna, Physica A **179**, 232 (1991).
12.14 R.N. Mantegna and H.E. Stanley, Nature **376**, 46 (1995).
12.15 S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge, Nature **381**, 767 (1996).
12.16 M. Potters, R. Cont, and J.-P. Bouchaud, Europhys. Lett. **41**, 239 (1998).
12.17 P. Gopikrishnan, M. Meyer, L.A.N. Amaral, and H.E. Stanley, Eur. Phys. J. B **3**, 139 (1998).
12.18 W. Li, Int. J. Bifurcat. Chaos **1**, 583 (1991).
12.19 R.N. Mantegna and H.E. Stanley, Nature **383**, 587 (1996).
12.20 Y. Liu, P. Cizeau, M. Meyer, C.-K. Peng, and H.E. Stanley, Physica A **245**, 437 (1997).
12.21 R. Cont, M. Potters, and J.-P. Bouchaud, in *Scale Invariance and Beyond*, edited by B. Dubrulle, F. Graner and D. Sornette (Springer, Berlin, 1997).
12.22 A. Arneodo, J.F. Muzy, and D. Sornette, Eur. Phys. J. B **2**, 277 (1998).
12.23 U. A. Müller, M.M. Dacorogna, R.B. Olsen, O.V. Pictet, M. Schwarz, and C. Morgenegg, J. Banking Finance **14**, 1189 (1995).
12.24 P. Cizeau, Y. Liu, M. Meyer, C.-K. Peng, and H.E. Stanley, Physica A **245**, 441 (1997).
12.25 J.-P. Bouchaud and R. Cont, Eur. Phys. J. B **6**, 543 (1998).
12.26 H. Takayasu, H. Miura, T. Hirabayashi, and K. Hamada, Physica A **184**, 127 (1992).
12.27 P. Bak, K. Chen, J. Scheinkman, and M. Woodford, Ric. Econ. **47**, 3 (1993).
12.28 D. Challet and Y.-C. Zhang, Physica A **256**, 514 (1998).
12.29 M. Lévy and S. Solomon, Int. J. Mod. Phys. C **7**, 595 (1996).
12.30 G. Caldarelli, M. Marsili, and Y.-C. Zhang, Europhys. Lett. **40**, 479 (1997).
12.31 M. Lévy, H. Lévy, and S. Solomon, J. Phys. I France **5**, 1087 (1995).
12.32 B.B. Mandelbrot, *Fractals and Scaling in Finance* (Springer, New York, 1997).
12.33 S. Maslov and Y.-C. Zhang, Physica **262**, 232 (1999).
12.34 P. Bak, M. Paczuski, and M. Shubik, Physica A **246**, 430 (1997).
12.35 A.H. Sato and H. Takayasu, Physica A **250**, 231 (1998).
12.36 D. Sornette and A. Johansen, Physica A **261**, 581 (1998).
12.37 D. Stauffer, Ann. Phys.-Berlin **7**, 529 (1998).
12.38 D. Stauffer and T.J.P. Penna, Physica A **256**, 284 (1998).
12.39 H. Takayasu, A.H. Sato, and M. Takayasu, Phys. Rev. Lett. **79**, 966 (1997).
12.40 T. Lux, J. Econ. Dyn. Control **22**, 1 (1997).
12.41 T. Lux, J. Econ. Behav. Organ. **33**, 143 (1998).
12.42 T. Lux and M. Marchesi, Nature **397**, 498 (1999).
12.43 J.-P. Bouchaud and D. Sornette, J. Phys. I France **4**, 863 (1994).
12.44 E. Aurell and S.I. Simdyankin, Int. J. Theor. Appl. Finance **1**, 1 (1998).
12.45 R. Baviera, M. Pasquini, M. Serva, and A. Vulpiani, Int. J. Theor. Appl. Finance **1**, 473 (1998).
12.46 S. Galluccio and Y.-C. Zhang, Phys. Rev. E **54**, R4516 (1996).
12.47 S. Galluccio, J.-P. Bouchaud, and M. Potters, Physica A **259**, 449 (1998).
12.48 M. Marsili, S. Maslov, and Y.-C. Zhang, Physica A **253**, 403 (1998).
12.49 D. Sornette, Physica A **256**, 251 (1998).
12.50 R.N. Mantegna and H.E. Stanley, Physica A **239**, 255 (1997).
12.51 J.D. Farmer, Market force, ecology, and evolution; [available from xxx.lanl.gov/abs/adap-org/9812005].
12.52 N. Vandewalle and M. Ausloos, Physica A **246**, 454 (1997).
12.53 M.H.R. Stanley, L.A.N. Amaral, S.V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M.A. Salinger, and H.E. Stanley, Nature **379**, 804 (1996).

12.54 L.A.N. Amaral, S.V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M.A. Salinger, H.E. Stanley, and M.H.R. Stanley, J. Phys. I France **7**, 621 (1997).
12.55 L.A.N. Amaral, S.V. Buldyrev, S. Havlin, M.A. Salinger, and H.E. Stanley, Phys. Rev. Lett. **80**, 1385 (1998).
12.56 H. Takayasu and K. Okuyama, Fractals **6**, 67 (1998).
12.57 Y. Lee, L.A.N. Amaral, D. Canning, M. Meyer, and H.E. Stanley, Phys. Rev. Lett. **81**, 3275 (1998).
12.58 E. Majorana, Scientia **36**, 58 (1942).
12.59 L.P. Kadanoff, Simulation **16**, 261 (1971).
12.60 E.W. Montroll and W.W. Badger, *Introduction to Quantitative Aspects of Social Phenomena* (Gordon and Breach, New York, 1974).
12.61 P.W. Anderson, J.K. Arrow, and D. Pines (eds.), *The Economy as an Evolving Complex System* (Addison-Wesley, Redwood City, 1988).
12.62 V. Pareto, *Cours d'Economie Politique*, (F. Rouge, Lausanne and Paris, 1897).
12.63 P. Lévy, *Calcul des Probabilités* (Gauthier-Villars, Paris, 1925).
12.64 B.B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).
12.65 H.E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, Oxford, 1971).
12.66 A. Einstein, Ann. Phys. **17**, 549 (1905).
12.67 N. Wiener, J. Math. Phys. **2**, 131 (1923).
12.68 B.B. Mandelbrot, J. Bus. **36**, 394 (1963).
12.69 B.V. Gnedenko and A.N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables* (Addison-Wesley, Cambridge, 1954).
12.70 E.F. Fama, J. Finance **25**, 383 (1970).
12.71 P.A. Samuelson, Industrial Management Rev. **6**, 41 (1965).
12.72 J.L. Doob, *Stochastic Processes* (Wiley, New York, 1953).
12.73 E.F. Fama, J. Finance **46**, 1575 (1991).
12.74 J.C. Hull, *Options, Futures, and Other Derivatives, 3rd edn.* (Prentice-Hall, Upper Saddle River, 1997).
12.75 D. Duffie and J. Pan, J. Deriv. [spring issue], 7 (1997).
12.76 Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.-K. Peng, and H.E. Stanley, Phys. Rev. E **60**, 1390 (1999).
12.77 Z. Palágyi and R N. Mantegna, Physica A **269**, 132 (1999).
12.78 R. Baviera, M. Pasquini, M. Serva, D. Vergni, and A. Vulpiani, Efficiency in foreign exchange markets (1999); [available from `xxx.lanl.gov/abs/cond-mat/9901225`].
12.79 M.M. Dacorogna, U.A. Müller, R.J. Nagler, R.B. Olsen, and O.V. Pictet, J. Int. Money Finance **12**, 413 (1993).
12.80 M. Pasquini and M. Serva, Physica A **269**, 140 (1999).
12.81 M. Raberto, E. Scalas, G. Cuniberti, and M. Riani, Physica A **269**, 148 (1999).
12.82 R.F. Engle, Econometrica **50**, 987 (1982).
12.83 T. Bollerslev, J. Economet. **31**, 307 (1986).
12.84 P.K. Clark, Econometrica **41**, 135 (1973).
12.85 R.C. Merton, J. Financ. Econ. **3**, 125 (1976).
12.86 R.N. Mantegna and H.E. Stanley, Phys. Rev. Lett. **73**, 2946 (1994).
12.87 T. Lux, Appl. Financ. Econ. **6**, 463 (1996).
12.88 A.L. Tucker, J. Bus. Econ. Statist. **10**, 73 (1992).
12.89 P. Gopikrishnan, V. Plerou, L.A.N. Amaral, M. Meyer, and H.E. Stanley, Phys. Rev. E **60**, 5305 (1999).
12.90 V. Plerou, P. Gopikrishnan, L.A.N. Amaral, M. Meyer, and H.E. Stanley, Phys. Rev. E **60**, 6519 (1999).
12.91 H. Markowitz, *Portfolio Selection: Efficient Diversification of Investment* (Wiley, New York, 1959).
12.92 R.N. Mantegna, Eur. Phys. J. B **11**, 193 (1999).
12.93 L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, Phys. Rev. Lett. **83**, 1468 (1999).
12.94 V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, and H.E. Stanley, Phys. Rev. Lett. **83**, 1471 (1999).
12.95 A.M. Sengupta and P.P. Mitra, Phys. Rev. E **60**, 3389 (1999).
12.96 P. Gopikrishnan, B. Rosenow, V. Plerou, and H.E. Stanley, Phys. Rev. E. Rapid Communications **64**, 035106-1 (2001).