# Dynamics of stock prices

Rosario Bartiromo*
*Istituto di Struttura della Materia del CNR, via Fosso del Cavaliere 100, I-00133 Roma, Italy*
*and Unitá INFM, Universitá di Roma Tre, via della Vasca Navale 84, 00146 Rome, Italy*
(Received 15 October 2003; published 15 June 2004)

We show that the dynamics of stock prices can be accurately described as a continuous time random walk with a time dependent diffusion coefficient. The time evolution of the diffusion coefficient can be derived from tick by tick databases provided the stock price is characterized in terms of a couple of values describing the best ask and the best bid. We are then led to a finding and, namely, that the transition rate of the random walk process is different from the frequency of transactions. Our results allow us to obtain a fast and reliable determination of the diffusion coefficient and precisely confirm that fat tails in the distribution of price variations are due to volatility fluctuations.

DOI: 10.1103/PhysRevE.69.067108 PACS number(s): 89.65.Gh, 05.40.Fb, 05.45.Tp, 89.75.Da

The idea that share prices are formed in stock markets as a result of a random walk process is deeply rooted in both economist's [1] and physicist's literature [2]. In practice it forms the basis for all methods used to price derivative financial products [3]. The key parameter in the pricing procedure is the asset volatility, closely related to the diffusion coefficient of the underlying random walk. Therefore a reliable and fast evaluation of this parameter is of paramount importance to monitor the risk involved in the trading activities.

Nowadays physicists are used to measuring directly the step size and the waiting time distribution characterizing random walks. This microscopic approach to diffusion yields the best time resolution and is very popular due to the availability of a variety of tools to perform atomic scale microscopy [4]. A microscopic picture of stock behavior is far more easy to obtain: every day most stock markets make freely available on their website a tick by tick database, where each transaction is registered [5]. Our aim is to show how to use them properly to evaluate diffusion and hence volatility.

Detailed analysis of the distribution of returns, as the variations of price logarithms $G_{\Delta t}(t)=\ln x(t+\Delta t)-\ln x(t) \approx \Delta x/x$ are referred to in the specialized literature, has been presented by a number of authors [6,7] showing marked deviation from a Maxwellian curve and the presence of fat tails at all time scales except very long ones. The question naturally arises of whether the observed distributions can still be interpreted as a result of a stationary statistical process. In his pioneering work on cotton price, Mandelbrot [8] put forward the idea that a description in terms of a stable Levy process was a good candidate to explain the data.

The analysis of large price databases has shown that the cumulative distribution representing these tails can be well fitted by a power law $P(x)=x^{-\alpha}$, as appropriate for a Levy distribution. However the exponent $\alpha$ turns out to be larger than 2 for both stocks [9], stock averages [10], bonds [7], and foreign exchange rates [11], with possibly the notable exception of spot commodity prices [12]. This rules out the possibility that the elementary distribution of returns has infinity variance, in contrast with the Levy hypothesis.

If the hypothesis of stationarity is abandoned, it is possible to account for the existence of fat tails provided the variance of the elementary distribution fluctuates on a time scale comparable to the time span required to fulfill the conditions for the validity of the central limit theorem in the wings of the distribution [7]. Market models which assume that the fractional standard deviation, also referred to as price volatility, is a stochastic variable are widely adopted in the economic literature.

From a physicist standpoint, these processes are best described as one-dimensional, continuous time, random walks [13] (CTRW). In these models the system makes a random jump $\delta$ and remains in the new position for a finite time $\tau$ before a new jump is made. If both the jump amplitude $\delta$ and the waiting time $\tau$ are statistically independent, the process is described by two distribution functions $f(\delta)$, which represents the probability density of jumps, and $\psi(\tau)$, the waiting time distribution, which gives the probability density of a pause $\tau$ between two successive steps.

It has been shown by Montroll and Schleisinger [14] that the CTRW propagator is governed by a generalized master equation which, for generic $f(\delta)$ and $\psi(\tau)$, has a non-Markovian character in both time and space domains. The results of standard random walk with a Gaussian propagator are recovered if $f(\delta)$ has finite second moment $\langle \delta^2 \rangle$ and $\psi(\tau)$ has finite first moment $\langle \tau \rangle$. In this case the diffusion coefficient is $D=\langle \delta^2 \rangle / 2\langle \tau \rangle$ and the price variation $\Delta x$ in the time interval $\Delta t$, once normalized to $\sqrt{2D\Delta t}$, should be normally distributed with unity standard deviation. In this paper we show that new insight on the dynamics of the process can be obtained by checking this basic conclusion against market data.

Differently from previous applications of CTRW to financial data [15,16], we do not make hypothesis about $f(\delta)$ and $\psi(\tau)$.We just assume that both $\langle \delta^2 \rangle$ and $\langle \tau \rangle$ are finite. The first assumption is justified by the observation that tails in the price distribution are outside the Levy regime. For the second assumption we can argue for the time being that the efficient market hypothesis [17] should rule out any non-

*Email address: bartiromo@fis.uniroma3.it

 

Markovian behavior in the time domain and therefore require the existence of a finite $\langle\tau\rangle$. We will come back later to this point once a proper definition of the waiting time in our problem will be in hand.

It is appropriate to note at this point that modern financial markets are intrinsically seasonal. Indeed typically trading activity is intense at the market opening, when the information from overnight events has to be taken care of, and before the closing, when traders need to adjust their position ahead of the night pause. Similarly on longer time scales, trading intensity fluctuates depending on the calendar of events such as, for example, management reporting and dividend distribution. Besides these structural cycles, additional variations come from the evolving perception of financial risks by trading agents which also has a cyclical behavior although with variable time period. The presence of different agents on the market with different aims and/or different perception of the risks involved in the trading activity makes it clear that both $\langle\delta^2\rangle$ and $\langle\tau\rangle$ have to be considered as time dependent quantities [7]. We will note this time dependence by an apex in the following of this paper.

For this work we use a tick by tick transaction database consisting of about three years of data of a future on the MIB average index (FIB) of the Milan stock exchange [18] and of 13 among the most traded shares on that market which we have selected so as to be representative of the different trading situations. We first analyze the distribution of returns to show that our database displays the typical behavior for this kind of assets. To ease comparison, we normalize for each stock the return $G_{\Delta t}(t)$ to its variance in the considered data set. In the upper panel of Fig. 1 we show the cumulative distribution function of returns of the FIB index on four time scales ranging from 15 min to more than 3 h. This figure shows the typical leptokurtic behavior reported in the literature, where the distributions stay below the Gaussian for normalized returns up to about 2 and fat tails appear above this value. The slow convergence toward the Gaussian with increasing time scale is also apparent. The straight line in the figure represents a power law with $\alpha=3$ and demonstrate that the experimental distributions are well outside the Levy stable regime.

In the lower panel of Fig. 1 we plot the cumulative distribution of 75 min minutes returns for all the 14 time series used in this study. Although the distributions remain very similar in their bulk, the $\alpha$ exponent which describes the tails can vary between 3 and 5 depending on the stock under consideration. This range of $\alpha$ is well representative of the values reported in the literature.

We move now to analyze the behavior of normalized price variation $\Delta x/\sqrt{2D\Delta t}=\Delta x/\sqrt{2\langle\delta^2\rangle_t\Delta t/\langle\tau\rangle_t}$. In first instance we compute $\langle\delta^2\rangle_t$ and $\langle\tau\rangle_t$ by considering each transaction as reported in the database. We focus our analysis on two time scales: 15 min, the minimum required to have a meaningful number of transactions, and 1 day, the maximum possible with our approach without making additional assumptions. Indeed price variations over a time interval longer than 1 day incorporate overnight changes which take place in the absence of trading activity.

We begin by evaluating the standard deviation $\sigma$ for all our 14 data series. The results are reported as full symbols in

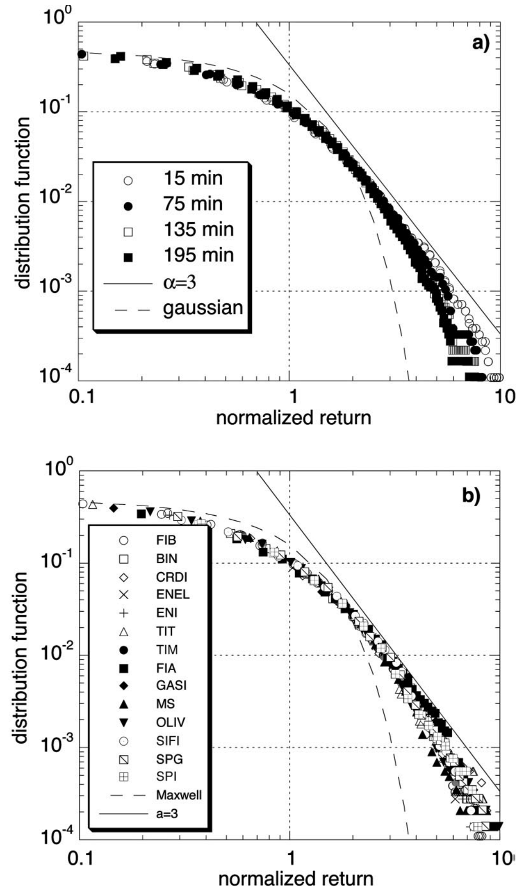


FIG. 1. (a) The cumulative distribution function of returns of the FIB index is shown for four time scales ranging from 15 to 195 min. On the abscissa the return is normalized to the standard deviation of each database, after subtracting the mean value. The positive and negative branches of the distribution are overlaid. The continuous line represents a power law with $\alpha=3$ while the dashed line is a normal distribution of unity variance. (b) Here we show the cumulative distribution of 75 min returns for all the 14 time series used in this study.

Fig. 2. We observe for both time scale values which are systematically below unity with average value of $0.47\pm0.03$ and $0.51\pm0.03$, respectively, for the short and the long time scale. Deviations from a Gaussian are also detected by the analysis of the kurtosis of the distributions which shows values systematically higher than 3, which is the Gaussian value. We obtain an average value of $3.77\pm0.1$ and $3.24\pm0.12$, respectively, for the short and the long time scale.

These findings suggest that the evaluation of the diffusion coefficient is biased and gives systematically high values. A careful consideration of the working mechanism of a stock market supports this indication. Indeed stock prices are the result of a continuous double auction whereby traders submit limit orders, where the less favorable price for the transaction is specified together with the number of shares to buy or sell. So called market orders also exist, where only the quantity is specified, which are therefore immediately executed at the best available price. It is also worth noting that prices are quantized and that the minimum price variation, the tick, can

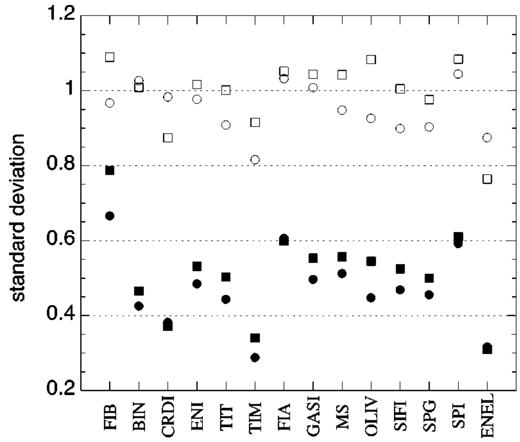


FIG. 2. The standard deviation $\sigma$ of normalized price variation is plotted for all the 14 series used in this study. Circles refer to 15 min data while squares refer to 1 day data. Full symbols are obtained by considering all transactions in the time series while open symbols show the results obtained when the system is described by the best bid–best ask prices. In this last case $\sigma$ is unity as expected from a CTRW description.

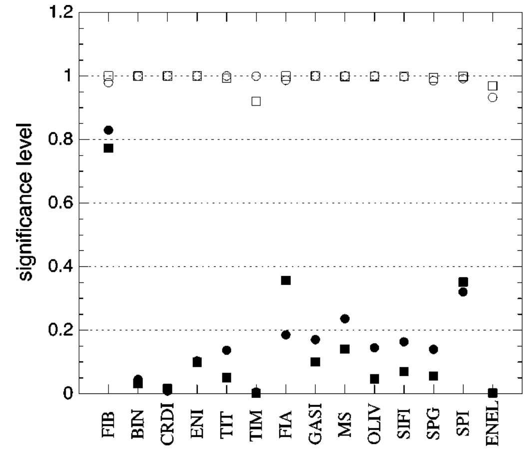


FIG. 3. The significance level of the Kolmogorov-Smirnov test. Circles refer to 15 min data while squares refer to 1 day data. Full symbols are obtained by considering all transactions in the time series while open symbols show the results obtained when the system is described by the best bid–best ask prices. The full symbols show a satisfactory significance level only for the FIB data while all the data set can pass the test when the open symbols are considered.

be different from stock to stock. The key device of a double auction is the order book where waiting limit orders are stored in two queues, the bids which are buy orders and the asks which are sell order. Therefore a double auction is not characterized by a single price but rather by the best bid, which is the price a potential seller would get, and the best ask, which is the price a potential buyer would need to pay in order to get the shares.

The double auction mechanism and its relevance to price formation is gaining an increasing attention in the physics literature [19]. Here we limit ourself to note that from a physicist point of view as long as the best bid and the best ask price remain constant in a double auction the state of the system is not changed. Any switch of the last contract price between a constant pair of best bid–best ask therefore does not represent an event, and the system has to be considered in a waiting state until this pair does not change.

Another way to look at this problem is to consider, for example, that once the last contract price makes a transition upward from the best bid to the best ask, it is more probable to observe a downward transition than one further step up. In fact a trader selling a single share can easily produce the step down while the whole amount of shares offered at the best ask has to be bought to produce a new step in the up direction. Therefore the time series of last price variation cannot be considered uncorrelated, the conditions for the validity of the central limit theorem are not satisfied and normal diffusion can be observed only over time intervals longer than the correlation time.

A practical way to overcome this difficulty is to identify a transition only when the best ask or the best bid changes and to quantify the step amplitude as the corresponding variation (usually corresponding to one tick for liquid stocks). A simple algorithm can easily identify most of the transitions from the price sequence in a transaction database (transitions are only missed when a new best ask or best bid is exposed but not touched by traders before reverting to the old value by an incoming limit order, a seldom occurrence in our experience). This leads to a different evaluation of both $\langle \delta^2 \rangle_t$ and $\langle \tau \rangle_t$.

We have therefore repeated our analysis using these new values in the computation of normalized returns. The results we obtained for $\sigma$ are reported again in Fig. 2 this time as open symbols. Now the values are well distributed around the unity with average value of $0.96 \pm 0.02$ and $1.01 \pm 0.03$, respectively, for the short and the long time scale. The Gaussian nature of the distribution is confirmed by the analysis of the kurtosis with average values of $2.93 \pm 0.04$ and $3.07 \pm 0.08$, respectively, for the short and the long time scale.

A more comprehensive and detailed comparison of the experimental distributions of normalized price variations

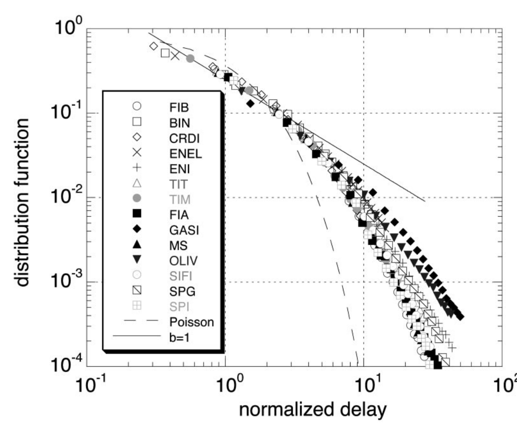


FIG. 4. This figure shows the cumulative distribution function of the waiting time between two successive transitions for all the 14 time series. On the abscissa we use the delay normalized to its mean value in the relative time series. The full line represents a power law with $\beta = 1$ while the dashed line is a Poisson distribution with unit average delay.

with a normal distribution of unitary variance was obtained by a Kolmogorov-Smirnov test on all 14 data sets. The outcome is illustrated in Fig. 3. As in the previous figures the full symbols show the results obtained by using all transactions. In this case only for the FIB index we obtain a significative probability to pass the test. This happen because the FIB is traded with a very small relative tick, less than $10^{-4}$, and therefore almost every transaction results in a transition. This is not the case for the other series whose tick ranges between $3\times10^{-3}$ and $10^{-3}$. None of them can pass satisfactorily the test.

The situation changes radically when a description in terms of transitions is adopted, as the open symbols in Fig. 3 clearly show. In this case the test is fully passed by all the data sets with a significance level not lower than 92%. This confirms the need to distinguish between transactions and transitions to better describe the dynamics of share prices.

We can now turn back to the waiting time distribution and analyze our database to verify the assumption we made about $\langle\tau\rangle$. In Fig. 4 we show the cumulative distribution function of the waiting time between two successive transitions for all the 14 time series. If we adopt a power law $\psi(\tau)=\tau^{-\beta}$ to fit the wing of the distributions in Fig. 4, we obtain $\beta$ ranging between 2 and 3. Since for all series we have $\beta>1$, these data confirm that $\langle\tau\rangle$ is finite as assumed in our previous analysis.

In summary we have shown that a continuous time random walk model is fully compatible with the observed fluctuations of stock prices once the time evolution of the variance of the step distribution and of the transition rate is taken into account. However, to achieve this result it is important that the system is characterized in terms of a couple of best ask–best bid prices and that a transition is identified only when this couple changes. We believe this is an important methodology remark to keep in mind whenever the dynamic of a stock market is analyzed by means of tick by tick data. We have shown that our method yields a fast and accurate determination of the diffusion coefficient thereby precisely confirming that fat tails in the distribution of price variations are due to volatility fluctuations.


[1] L. Bachelier, Ann. Sci. Ec. Normale Super. **17**, 21 (1900).

[2] R. N. Mantegna and H. E. Stanley, Nature (London) **376**, 46 (1995).

[3] F. Black and M. Scholes, J. Polit. Econ. **3**, 637 (1973).

[4] D. C. Senft, and G. Ehrlich, Phys. Rev. Lett. **74**, 294 (1995).

[5] M. Dacorogna, R. Gençay, U. Müller, R. Olsen, and O. Pictet, *An Introduction to High-Frequency Finance* (Academic Press, London, 2001).

[6] R. N. Mantegna and H. E. *Stanley, An Introduction to Econophysics* (Cambridge University Press, Cambridge, 1999).

[7] J. P. Bouchaud and M. Potters, *Theory of Financial Risks* (Cambridge University Press, Cambridge, 2000).

[8] B. B. Mandelbrot, J. Business **36**, 294 (1963).

[9] V. Plerou, P. Gopikrishanan, L. A. Nunes Amaral, M. Maier, and H. E. Stanley, Phys. Rev. E **60**, 6519 (1999).

[10] P. Gopikrishanan, V. Plerou, L. A. Nunes Amaral, M. Maier, and H. E. Stanley, Phys. Rev. E **60**, 5305 (1999).

[11] U.A. Müller, M.M. Dacorogna, R.B. Olsen, O.V. Pictet, M. Schwarz, and C. Morgenegg, J. Banking Finance **14**, 1189 (1995).

[12] K. Matia, L. A. Nunes Amaral, S. P. Goodwin, and H. E. Stanley, Phys. Rev. E **66**, 045103(R) (2002).

[13] E. W. Montroll and G. H. Weiss, J. Math. Phys. **6**, 167 (1965).

[14] E. W. Montroll and M. F. Schleisinger, in *Studies in Statistical Mechanics*, edited by J. L. Lebowitz and E. W. Montroll, (North Holland, Amsterdam, 1983), Vol. 11, p. 5.

[15] E. Scalas, R. Gorenflo, and F. Mainardi, Physica A **284**, 376 (2000).

[16] J. Masoliver, M. Montero, and G. H. Weiss, Phys. Rev. E **67**, 021112 (2003).

[17] E. F. Fama, J. Business **38**, 34 (1965).

[18] See http://www.borsaitalia.it for the data used in this work.

[19] M. Daniels, J. D. Farmer, L. Gillemot, G. Iori, and D. E. Smith, Phys. Rev. Lett. **90**, 108102 (2003).