Physica A 392 (2013) 1638–1647

Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Maximum entropy distribution of stock price fluctuations

Rosario Bartiromo *

Istituto di Struttura della Materia del CNR, via Fosso del Cavaliere 100, 00133 Roma, Italy
Dipartimento di Fisica, Università di Roma Tre, via della Vasca Navale 84, 00146 Roma, Italy

ARTICLE INFO

Article history:
Received 25 May 2012
Received in revised form 8 November 2012
Available online 27 November 2012

Keywords:
Econophysics
Financial markets
Time series analysis
Information theory
Entropy
Derivative pricing

ABSTRACT

In this paper we propose to use the principle of absence of arbitrage opportunities in its entropic interpretation to obtain the distribution of stock price fluctuations by maximizing its information entropy. We show that this approach leads to a physical description of the underlying dynamics as a random walk characterized by a stochastic diffusion coefficient and constrained to a given value of the expected volatility, in this way taking into account the information provided by the existence of an option market. The model is validated by a comprehensive comparison with observed distributions of both price return and diffusion coefficient. Expected volatility is the only parameter in the model and can be obtained by analysing option prices. We give an analytic formulation of the probability density function for price returns which can be used to extract expected volatility from stock option data.

1. Introduction

The behaviour of stock prices has attracted considerable attention by the physics community in recent years [1–3]. More than by its practical interest for asset allocation and risk management, physicists are particularly intrigued by two features of the distribution of price returns $x(t) = \log(p(t)/p(t - \Delta t))$, where Δt is the observation time and $p(t)$ the price value at time t. First, it is markedly non-Gaussian, with fat tails enhancing the probability of large fluctuations by many orders of magnitude [4]. Second, when returns are normalized to their standard deviation, different stocks and different markets show very similar probability distributions or in other words the return distribution scales to a common shape [4,5], as we will document again in the following. Moreover this universal distribution shows a mild dependence upon the observation time Δt as long as it remains shorter than 16 trading days [6]. The challenge to find a simple mechanism to explain this neat behaviour has not yet been met.

A description of the observed phenomenology in terms of a stable Levy process was first put forward by Mandelbrot [7] and analysed in great detail in subsequent years by many authors over a large body of market data. It requires that the cumulative distribution of large returns is well described by a power law, in reasonable agreement with observations [5,6]. However the exponent in this power law should be higher than 2 to be compatible with a diverging variance as required by the Levy hypothesis, a feature which is not confirmed by observations. We note in passing that a diverging variance of price returns would make very difficult, if not impossible, the task of quantifying financial risks.

An alternative approach to account for the existence of fat tails assumes that the price distribution is not stationary but depends on market conditions [8,9]. More specifically, a Gaussian distribution is expected for a stationary market, since the stock price would depend on the stochastic behaviour of a large number of market agents and, in force of the central limit theorem, this would produce a Gaussian. The standard deviation of this elementary distribution would then quantify the risk of keeping the stock in portfolio and thus should in principle be a time dependent quantity. This recalls a class of

* Correspondence to: Dipartimento di Fisica, Università di Roma Tre, via della Vasca Navale 84, 00146 Roma, Italy. Tel.: +39 0657337217.
E-mail address: bartiromo@fis.uniroma3.it.

0378-4371/$ – see front matter
doi:10.1016/j.physa.2012.11.048

physical models, collectively known as superstatistics [10], whereby fluctuations of an intensive parameter are postulated to describe the turbulent behaviour of non-equilibrium systems [11,12].

In the following in this paper we will refer to the standard deviation of price return distribution as the price volatility. Stochastic volatility models [13] assume either a specific volatility distribution [14] or a parametric form for an auxiliary process which regulates the instantaneous value for the width of the elementary Gaussian distribution [15]. In econometrics autoregressive models, known by the acronym ARCH [16] and its numerous modifications [17–20], were developed to forecast volatilities and, depending on the particular version adopted, have been shown able to reproduce many of its observed characteristics such as long range memory [21] and leverage effects [22]. Optimal values for the model parameters can be obtained by fitting the observed distribution of price returns [23,24] or, as recently done by Gerig et al. [25], by fitting the observed distribution of daily volatility. Although these models are able to follow the time evolution of volatility and can be usefully adopted for a variety of tasks related to risk assessment and derivative contract pricing, they remain phenomenological in nature and do not tell us very much about the mechanism which leads to the observed price fluctuations.

In this paper we will show that a useful expression for price return distribution can be obtained with an appropriate exploitation of the principle of maximum entropy without any detailed assumption about the volatility dynamics. We will be led by a physics based description of market mechanisms to a double stochastic model suitable to reproduce scaling and universality of price fluctuations. This will enable us to compute the entropy of the process. Subsequently we will assume that the mere observation of price fluctuations does not provide any new information, as implied by the principle of absence of arbitrage opportunities. This allows maximizing the process entropy which we exploit to find the marginal volatility distribution and finally the price fluctuation distribution. The final result depends critically on the public information available to market participants and we will use market data to discriminate among different plausible options. This analysis will show that the best physical description of price fluctuation dynamics is a random walk characterized by a stochastic diffusion coefficient and constrained to a given value of the expected volatility, in this way taking into account the information provided by the existence of an option market.

The paper is structured as it follows: we will first describe in Section 2 how a stock market works and we will introduce the concept of an ideally liquid stock to simplify its physical description. The basic formalism of a double stochastic model will be given in Section 3 while in Section 4 we will calculate its information entropy and discuss the procedure and the constraints used to maximize its value. An extensive comparison with market data for daily price returns will be presented in Section 5 and in the following Section 6 use will be made of tick data to evaluate diffusion coefficients to compare with the predictions of our models. Section 7 will be devoted to the issue of stationarity and its influence on the tails of the distribution. In the last section we will discuss our conclusions.

2. Stock market operation

Stock prices are set in a continuous double auction [26] whereby trader orders that are not immediately executed at the best available price are stored in the order book in two queues, the bids which are buy orders and the asks which are sell orders. Prices are quantized and the minimum price variation, the tick τ, can be different from stock to stock. 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. In a previous work [27] we have shown that to describe the stochastic dynamics of stock prices as a continuous time random walk the system has to be considered in a waiting state as long as both the best bid and the best ask prices remain constant. A simple algorithm can then easily identify a transition from the price sequence in a transaction database when the mid-price between the best ask and the best bid changes and quantify the step amplitude as the corresponding variation (usually corresponding to one tick for liquid stocks).

To simplify the understanding of our analysis we introduce the concept of an ideally liquid stock [28]. We shall say that a stock is ideally liquid when all levels in its order book are always filled and its mid-price can only change by the tick size. Then the stochastic dynamics of an ideally liquid stock is that of a binomial random walk characterized by the number N of transitions recorded during the observation time Δt and the probability η to observe an upward transition. We take η equal to $1/2$ neglecting the small variations allowed by the efficient market hypothesis. Under these assumptions the price return should perform a random walk with diffusion coefficient given by $D = \frac{\tau^2}{p^2} f$ where f is the transition frequency. Price return volatility on the scale Δt is therefore given by $w = \sqrt{D\Delta t} = \sqrt{N\left(\frac{\tau}{p}\right)^2}$ provided the observation time Δt is long enough to allow a number of transitions higher than about 10. If w does not change with time the return distribution is normal with zero mean and variance equal to w^2. However the transition frequency is observed to change both on a day to day and on an intraday basis, therefore both volatility and diffusion must be time dependent quantities in this approach.

This analysis of market mechanisms suggests that traders may adapt to changing risk perception by varying the rate at which they trade and therefore mainly by means of the transition frequency f. In extreme situations they could also open the bid–ask spread to more than one tick but in both circumstances it appears that from a physicist's standpoint the diffusion coefficient would be the dynamical variable that controls the volatility value w.