Physics Letters A 380 (2016) 2501–2512

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Universal behaviour in the stock market: Time dynamics of the electronic orderbook

Ayşe Kızılersü [a,*], Markus Kreer [b,1], Anthony W. Thomas [c], Michael Feindt [d,2]

[a] Special Research Centre for the Subatomic Structure of Matter (CSSM), Department of Physics, School of Chemistry and Physics, Adelaide University, 5005, Australia
[b] phi-t products & services, Karlsruher Strasse 88, 76139 Karlsruhe, Germany
[c] CoEPP and CSSM, Department of Physics, Adelaide University, SA 5005, Australia
[d] Blue Yonder GmbH, Ohiostraße 8, 756139 Karlsruhe, Germany

ARTICLE INFO

Article history:
Received 24 November 2015
Received in revised form 20 May 2016
Accepted 20 May 2016
Available online 1 June 2016
Communicated by C.R. Doering

Keywords:
Econophysics
Stock market analysis
Statistical test
Financial physics
Orderbook dynamics
Stochastic time analysis

ABSTRACT

A consequence of the digital revolution is that share trading at the stock exchange takes place via electronic order books which are accessed by traders and investors via the internet. Our empirical findings of the London Stock Exchange demonstrate that once ultra-high frequency manipulation on time scales less than around ten milliseconds is excluded, all relevant changes in the order book happen with time differences that are randomly distributed and well described by a left-truncated Weibull distribution with universal shape parameter (independent of time and same for all stocks). The universal shape parameter corresponds to maximum entropy of the distribution.

1. Introduction

In a functioning stock market, shares are exchanged for a certain amount of money at specific times. There are traders who want to buy and those who wish to sell. The electronic order book (EOB) enables this trading to occur in a regulated environment. Understanding the probability distributions for buying and selling stocks is of considerable importance, not just because of the potential for financial gain but also in assessing the stability of the market. An accurate prediction of the occurrence of the catastrophic movements is crucial to being prepared for them and minimising their impact. We make a crucial first step in this direction by establishing that the London Stock Exchange (LSE) exhibits a remarkable degree of universality in the timing of orders sent to the EOB, provided the intervals between orders is greater than typically 10 milliseconds.

Naturally the sellers want to sell for the highest possible price and they will indicate what they are willing to accept in the EOB by setting the "ask price", along with the number of shares they wish to sell. Likewise, the buyers want to buy for the lowest possible price and they will indicate the price they are prepared to pay in the EOB by the "bid price", along with the quantity they wish to purchase. The intention of buying or selling is termed a limit order (LO) and consequently the EOB consists of Ask-LOs and Bid-LOs, the Ask-prices are usually higher than the Bid-prices; otherwise the trader could purchase the stock at the lower Ask-price and sell immediately afterwards at the higher Bid-price and gain risk-free profit (arbitrage). The highest Bid-price is called "Best Bid" and the lowest Ask-price correspondingly "Best Ask". The midprice is the arithmetic mean of Best Bid and Best Ask. The midprice changes if and only if either Best Bid or Best Ask price changes (note these cannot change at the same time).

There are also market participants who are willing to buy or sell shares for the current "best price". These orders are called market orders (MOs) and they are matched with LOs. Thus the Best Bid or Best Ask will be cancelled by the trade (once the best price is exhausted the next best price will be taken and so on). This mechanism of trading is the so called "continuous double-auction" because trading takes place continuously within the trading session of the stock exchange and auctions happen both for buying and selling.

* Corresponding author.
E-mail address: ayse.kizilersu@adelaide.edu.au (A. Kızılersü).
[1] Present addresses: CAMPUSERVICE GmbH, Servicegesellschaft der Johann Wolfgang Goethe-Universität Frankfurt, 60323, Frankfurt am Main, Germany.
[2] University of Karlsruhe (on leave of absence from KIT), Germany.

http://dx.doi.org/10.1016/j.physleta.2016.05.035

Table 1
Passing rates for different distribution hypothesis for arrival times and some selected stocks with various truncations points τ_L and sample sizes N.

Ticker	τ_L [ms]	Weibull		log-normal		Pareto	
		$N=30$	$N=100$	$N=30$	$N=100$	$N=30$	$N=100$
	1	53%	13%	64%	7%	5%	0%
RIOTINTO	11	90%	81%	86%	54%	28%	2%
	101	**93**%	**89**%	**83**%	**62**%	**54**%	**11**%
	1	60%	19%	67%	8%	5%	0%
VOD	11	89%	78%	85%	52%	9%	0%
	101	**93**%	**89**%	**90**%	**74**%	**25**%	**1**%
	1	65%	25%	59%	7%	3%	0%
RRLN	11	81%	58%	71%	27%	6%	0%
	101	**92**%	**88**%	**87**%	**65**%	**9**%	**0**%

The Stock Exchange sends all EOB changes to data vendors (e.g. Reuters, Bloomberg, Morningstar, etc.), who in turn forward the information to the end user as a sequence of strings. A typical string contains an identification number, the name ticker of the stock company (e.g., VOD for the UK company Vodafone), type of order (bid or ask), limit price (ask, bid), quantity of shares to be traded, limit order status information ("addition" for a newly arrived orders, "cancellation" for cancelled orders or "modification" for changes to an existing order) and finally a trade indicator stating if the order has been cancelled by a trade (i.e., whether a market order has matched a limit order) or not.

For each stock, the times at which the limit orders are entered into the EOB are known as LO arrival times (t^a), the times at which the LO are cancelled are known as LO cancellation times (t^c) and the times at which market orders are entered into the EOB are known as the MO arrival times (t^{mo}).

The timing and direction of the next move in a stock price is something that analysts try to predict, especially since high-frequency "real time" trading became possible. In this paper we quantify the behaviour of

- The difference between consecutive LO arrival times, $x_i^a = t_i^a - t_{i-1}^a$
- The difference between consecutive LO cancellation times, $x_i^c = t_i^c - t_{i-1}^c$
- The difference between LO arrival and cancellation times for an individual entry, life times, $x_i^L = t_i^c - t_i^a$
- The difference between consecutive MO arrival times, $x_i^{mo} = t_i^{mo} - t_{i-1}^{mo}$

where $i = 1, 2, \cdots, N^{(EOB)}$ and $N^{(EOB)}$ is the number of EOB entries.

In this double auction EOB environment, limit orders appear to arrive and depart in a random fashion with observed "Poisson-type clumping" effects. We find that this "clumping" of time differences can be best described by a two-parameter Weibull distribution. The Weibull distribution [1] is widely used in reliability engineering and life data analysis, as it only considers positive random variables, is a minimum stable distribution [2,3], has a monotonic hazard rate and is a relatively simple function (see below) specified by two parameters, α and β. The former can be viewed as a scale (or inverse activity) and the latter as a shape parameter. Although the Weibull distributions for stocks has previously been considered by others [4–8], our analysis is the first to establish in statistical rigour that it describes the probability distributions of time differences in the EOB. Indeed, for a broad selection of stocks on the LSE, we establish that it describes the time differences between stock market orders far better than other distributions such as Pareto, log-normal etc. In order to quantify this we give the passing rates in Table 1 for some selected stocks (RIOTINTO, VOD, RRLN) at three different left-truncation points ($\tau_L = 1, 11, 101$ ms) for $N = 30, 100$ and plot the CDF and PDF comparisons of these distributions for RIOTINTO in Fig. 1 for ($\tau_L = 11, 101, 1001$ ms). Although the consensus among the practitioners and theoreticians is to describe the tail of the distribution is power-law, namely Pareto distribution [9], Table 1 and Fig. 1 show that in fact the Weibull distribution out-perform the power law for the EOB data. Note that the left-truncation procedure leaves the parameters of the one-point distribution function of the underlying stochastic process unchanged and just filters out the ultra-high frequency data. Possible autocorrelations of the stochastic process have not been investigated in the presented analysis.

The information content of a positive random variable, x, can be characterised by the continuous entropy (also known as the differential entropy) [10,11]

$$H \equiv -\int f(x) \log f(x) dx, \tag{1}$$

where $f(x)$ is the probability density function. The continuous entropy has the property that the more spread out the distribution is, the greater its value. For the Weibull distribution the maximum entropy depends only on the shape parameter, β, and is independent of the activity parameter, α. This means that if the behaviour of the trades for stocks self-organises to a state of maximum entropy, they can have wildly different activity but keep the same shape (that is a universal value of β).

Our hypothesis is that the distribution of time differences in the EOB is described by a maximum entropy Weibull distribution whose activity varies with the stock, time of day and which of the differences is being measured (e.g., the LO time of arrival difference).

1.1. Weibull distribution explains time differences between orders

We have studied the time differences, x, of LO arrival times, LO cancellation times, LO lifetimes and MO arrival times from the EOB for seven stocks (RIOTINTO, BARC, VOD, RRLN, ABFLN, SSELN, YELLLN) on the LSE over a four month period in 2010. The times in the EOB are recorded to millisecond precision and some of the distributors of this data keep the chronological order but group data into 10 millisecond blocks. Further investigation shows that there are many more time differences counted between zero and ten milliseconds than one would expect if the differences were described by a Weibull distribution. These orders are generated by ultra-high-frequency trading algorithms operating in the range of microseconds by computers situated in close proximity to the stock exchange building [12] (i.e. up to 1000 m nearby). The rapid addition and removal of LOs is thought to be an attempt to manipulate the stock market. We find that if these fast, frequent short term orders are included the distribution of time differences does not