World Scientific
www.worldscientific.com

THE KEY ROLE OF LIQUIDITY FLUCTUATIONS IN DETERMINING LARGE PRICE CHANGES

FABRIZIO LILLO

INFM Palermo Unit and Dipartimento di Fisica e Tecnologie Relative, Università di Palermo
viale delle Scienze, Edificio 18, I-90128, Palermo, Italy

J. DOYNE FARMER

Santa Fe Institute
1399 Hyde Park Road, Santa Fe, NM 87501, USA

Recent empirical analyses have shown that liquidity fluctuations are important for under-standing large price changes of financial assets. These liquidity fluctuations are quantified by gaps in the order book, corresponding to blocks of adjacent price levels containing no quotes. Here we study the statistical properties of the state of the limit order book for 16 stocks traded at the London Stock Exchange (LSE). We show that the time series of the first three gaps are characterized by fat tails in the probability distribution and are described by long memory processes.

Keywords: Financial markets; long-memory processes; power-law distributions.

1. Introduction

One of the best known statistical regularities of financial time series is the fact that the empirical distribution of asset price changes is fat tailed, i.e. there is a higher probability of extreme events than in a Gaussian distribution [1–3]. Moreover, there are strong indications that the part of the distribution describing large price changes is a power-law [4]. Many mechanisms explaining this empirical fact have been suggested, including subordinated stochastic processes [5] and volume fluctuations [6]. By investigating the fine details of the process of price formation in the London Stock Exchange, we have recently shown that large price fluctuations are driven by liquidity fluctuations, i.e. variations in the markets ability to absorb new orders [7]. We have shown that even for the most liquid stocks there can be substantial gaps in the limit order book, corresponding to a block of adjacent price levels containing no quotes (see below for a definition). When such a gap exists next to the best price, a new order can remove the best quote, triggering a large price change. Thus,
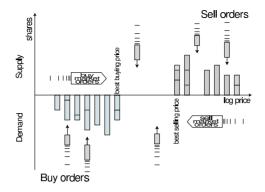
Fig. 1. A schematic illustration of the continuous double auction mechanism. We adopt the arbitrary convention that buy orders are negative and sell orders are positive.

the distribution of large price changes merely reflects the distribution of gaps in the limit order book. This is a finite size effect, caused by the granularity of order flow: in a market where participants place many small orders uniformly across prices, such large price fluctuations would not happen.

In this paper we investigate the statistical properties of the first three gaps of 16 stocks traded at the LSE. Specifically we investigate whether the fat tail properties of the first gap discovered in [7] are shared by other gaps. We then investigate the time correlation properties of gaps and the synchronous cross correlation between the size of different gaps. The analysis shows that the liquidity of the book as described by the multidimensional stochastic process of the gap sizes has interesting properties, such as power-law distribution, long memory and synchronous behavior.

In order to have a representative sample of high volume stocks we select 16 companies traded on the London Stock Exchange (LSE) in the 4-year period 1999-2002. The stocks we analyzed are Astrazeneca (AZN), Baa (BAA), BHP Billiton (BLT), Boots Group (BOOT), British Sky Broadcasting Group (BSY), Diageo (DGE), Gus (GUS), Hilton Group (HG.), Lloyds Tsb Group (LLOY), Prudential (PRU), Pearson (PSON), Rio Tinto (RIO), Rentokil Initial (RTO), Reuters Group (RTR), Sainsbury (SBRY), Shell Transport & Trading Co. (SHEL). These stocks were selected because they have high volume and they are all continuously traded during the full period. The number of transactions vary from $226,000$ for BAA to $723,000$ for LLOY. The total number of considered transactions is 6.5 million.

2.   Market Mechanism

To understand our results it is essential that the reader understand the double continuous auction, which is the standard mechanism for price formation in most modern financial markets. Agents can place different types of orders, which can be grouped into two categories: Impatient traders submit *market orders*, which are requests to buy or sell a given number of shares immediately at the best available price. More patient traders submit *limit orders*, or *quotes* which also state a limit price π, corresponding to the worst allowable price for the transaction. Limit orders

often fail to result in an immediate transaction, and are stored in a queue called the *limit order book*. Buy limit orders are called *bids*, and sell limit orders are called *offers*. At any given time there is a best (lowest) offer to sell with price $a(t)$, and a best (highest) bid to buy with price $b(t)$. These are also called the *best prices*. The price gap between them is called the *spread* $s(t) = a(t) - b(t)$. Prices are not continuous, but rather change in discrete quanta called *ticks*. The number of shares in an order is called either its *size* or its *volume*. As market orders arrive they are matched against limit orders of the opposite sign in order of first price and then arrival time, as shown in Fig. 1. A high density of limit orders per price results in high *liquidity* for market orders, i.e., it implies a small movement in the best price when a market order is placed. When a market order arrives it can cause changes in the best prices. This is called *market impact* or *price impact*. Buy market orders can increase the midprice $m(t) = (a(t) + b(t))/2$, and sell orders can decrease it. Price changes are typically characterized as *returns* $\Delta p_\tau(t) = \log m(t) - \log m(t-\tau)$.

3.   Importance of the Granularity of Limit Order Book

We investigate the unconditional probability density function of price changes at the level of individual transactions. To be more precise, for each transaction at time t we compute the price return, defined as $\Delta p(t) = \log(m(t+\epsilon)/m(t-\epsilon))$, where m is the midprice and ϵ is an infinitesimal time. The quantity Δp is termed the *price impact* and its average and conditional properties have been extensively investigated recently (see for example [8,9]). We have investigated the unconditional statistical properties of the price impact [7]. In Fig. 2 we plot the unconditional cumulative distribution of individual price changes for the stock AZN. The empirical distribution is consistent with a power-law distribution, i.e a probability distribution decaying asymptotically as $P(x > z) \sim z^{-\alpha}$, where α is called the tail index. A least squares fit on the tail of the distributions gives a value of the tail index equal to $\alpha = 2.19$ for the positive tail and $\alpha = 2.17$ for the negative tail. These results indicates that the stochastic process of the price impact has finite variance. At first sight the large fluctuations of price impact are puzzling. How can price changes due to individual transactions fluctuate so much?

A possible explanation is the role of the volume. Large individual price changes could be triggered by large market orders that cause transactions with many limit orders present in the book at different price levels. The effect of such a large market order would be a large price change. We have shown [7] that this is not the most frequent scenario. Specifically, the volume of the market order is important in determining whether the midprice changes, but when the midprice changes, the price return induced by the transaction is almost independent of the market order size. Thus the role of the volume of the market order in determining the individual price change is marginal. In other words, at the level of individual transactions, the impact function is not deterministic and the fluctuations of price impact are very large. These results show that the fluctuations in the state of the book have a key role in determining price changes. But how can small volume transactions create large price changes?

We have investigated this problem by studying the state of the limit order book before a individual market order was able to trigger a large price change. We have
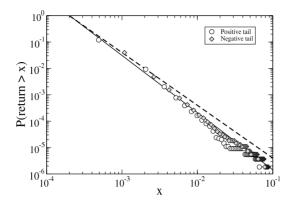
Fig. 2. Distribution of large price returns triggered by individual transactions (price impact). The continuous line is the best fit for the negative tail, whereas for illustrative purposes the dashed line indicates the power law behavior separating the regime with finite variance from the Lévy regime (infinite variance). The data refer to the stock Astrazeneca (AZN).

found [7] that an important way of describing the state of the book at a given time is through the properties of the *gaps*. To understand what we mean by a gap, let us consider the limit orders on the sell side of the book. By definition, the lowest price where shares are offered for sale is the best ask $\pi_0(t) \equiv a(t)$. The next price level occupied by (at least) one sell limit order will be denoted by $\pi_1(t) > \pi_0(t)$. In general the nth price level inside the book occupied by at least one limit order is $\pi_n(t)$, $n = 1, 2, ... N_s(t)$, where $N_s(t) + 1$ is the number of occupied offer price levels at time t. The first gap is defined as $g_1(t) = \log \pi_1(t) - \log a(t)$, and the nth order gap is defined as $g_n(t) = \log \pi_n(t) - \log \pi_{n-1}(t)$. Likewise the gaps for the side of the book occupied by buy limit orders is $g_{-n}(t) = \log \pi_{n-1}(t) - \log \pi_n(t)$ and $\pi_0(t) = b(t)$. Thus g_i is the ith gap starting from the best prices on the bid side for $i < 0$ and on the offer side for $i > 0$.

The size of the gaps (and the number of occupied price levels) present in the limit order book at a given time are a measure of the liquidity of the asset at that time. In fact a book with few occupied price levels and large gaps, i.e. a *sparse book*, is an indication of a lack of liquidity and reluctance to trade the asset by the investors. On the other hand when many orders are present and the gaps are small (*a dense book*) the amount of liquidity of the asset is high. The empirical analysis shows that the first gap size is a highly fluctuating quantity. Specifically it has been shown that the first gap size distribution $p(g_1)$ and $p(g_{-1})$ are well approximated in the tail by a power-law distribution with a small exponent. This implies that it is not uncommon that the best available limit price beside the best is very far from the best price. Although this is a clear indication of liquidity crisis, we still have not determined what causes these crises. We have shown that the first gap distribution is a key quantity to understand large individual price changes [7]. In fact we have shown that the first gap distribution is very similar to the distribution of nonvanishing price changes. The reason for that is that investors tend to place market orders of size smaller or equal to the size of the order at the best available
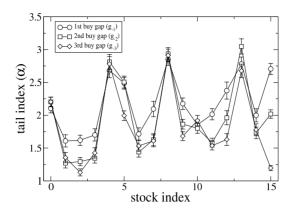
Fig. 3. Hill estimator of the tail index α of the tail of the distribution of the size of the first three gaps on the bid size of the limit order book (g_{-1}, g_{-2}, and g_{-3}). On the x axis we sort alphabetically the 16 stocks under investigation.

price (best bid or best ask). Thus when the size of the market order is smaller than the volume at the corresponding best, the midprice does not change and the price impact is zero. When the volume of the market order is equal to the volume at the best the midprice changes of a quantity $m(t + \epsilon) - m(t - \epsilon) = g_{\pm1}(t)/2$, where the plus (minus) sign refers to buy (sell) market orders. This argument shows why the distribution of first gap and individual price change are closely related.

In conclusion, our previous analysis has shown that the most important factor determining large price changes is not the transaction volume. We have shown instead that a liquidity crisis, characterized by a sparse state of the book, causes a situation in which a small market order can trigger a large price change. Even if we still do not have an explanation for what creates the conditions for the sparseness of the book, we believe that this is an important factor determining the dynamics of asset prices. In this sense it is important to investigate in more detail the statistical properties of the granularity of the limit order book. In this paper we provide a first empirical analysis of these properties.

A first remark to the above discussed explanation of the large price changes in terms of the fat tail of the first gap distribution concerns whether the first gap is in some sense special or whether the whole limit order book (i.e. also the second, third, etc. gaps) is characterized by granularity and large fluctuations. To this end we analyzed the probability distribution of the first three gaps both of sell and of buy limit orders. Specifically we compute the empirical pdf of gap size and we characterize the tail of the distribution using the Hill estimator [10]. This is a semi-parametric statistical estimator of the tail index α of a power-law distribution. Figure 3 shows the Hill estimator of the tail index α of the first three gaps of the buy side of the limit order book of the 16 stocks investigated (similar results are observed for gaps on the sell side). Figure 3 shows that the asymptotic statistical properties of the distribution of the first three gaps of a given stock are quite similar. This result supports the idea that the whole book follows a time dynamic in which the balance between the flows of limit orders, market orders and cancellations makes
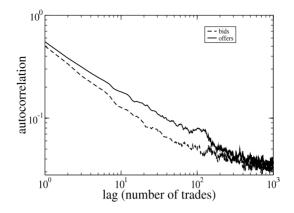
Fig. 4. Autocorrelation function of the first gap size for bids ($g_{-1}(t)$) and offers ($g_1(t)$) in a log-log plot. The data refer to Astrazeneca.

the book more or less granular (or sparse). The error bars given in Fig. 3 are 95% confidence intervals of the estimator, under the assumption that the data are IID. Unfortunately, these data are far from IID. The real error bars are certainly larger than these, due to the long-memory property of the data discussed below. It is difficult to assign proper error bars for long-memory processes, and doing so is beyond the scope of this paper. However, it is worth noting that the variation of the tail index across different stocks is much larger than the variation of the tail index of different gaps in the same stock. It is our belief that the variations between the tail exponents of gaps are not statistically significant.

The idea that the sparseness of the limit order book is not concentrated in "space", i.e. it is a property shared by the whole book, suggests that it is worth investigating whether the sparseness of the book is also persistent in time. More specifically, if the book is sparse at a given instant of time, will the sparseness persist in time, or is a sparse book an event that is strongly localized in time? To answer this question we have studied the time correlation properties of gap size. Figure 4 shows the autocorrelation function of the first buy and sell gap size of AZN in double logarithmic scale. The figure indicates that the first gap size is a persistent quantity. The autocorrelation function is well fit by a power-law $\rho(\tau) \sim \tau^{-\beta}$ with $\beta = 0.24$ for the first gap of buy limit orders (bids) and $\beta = 0.34$ for the first gap of sell limit orders (offers). These values are consistent with long memory processes characterized by an Hurst exponent $H = 1 - \beta/2$ larger than $1/2$. Long memory processes [11] are stochastic processes characterized by the lack of a typical time scale and they seems to provide a good description of several financial quantities such as the volatility [14] and the sign of orders [12,13]. It is known that the power-law fit of the autocorrelation function is not a good statistical estimator of the Hurst exponent. We have thus measured the Hurst exponent of the first three gaps for bids and offers by making use of the Detrended Fluctuation Analysis (DFA) [15] which is known to be a good estimator of long memory dependence. The result for AZN is $H(g_{-1}) = 0.75$, $H(g_{-2}) = 0.76$, $H(g_{-3}) = 0.76$ for gaps in bids and $H(g_1) = 0.78$,

Table 1. Correlation coefficient of the first three gap size g_i on the buy side ($i = -3, -2, -1$) and on the sell side ($i = 1, 2, 3$) of the limit order book. The data shown refer to the stock Astrazeneca.

ρ	g_{-3}	g_{-2}	g_{-1}	g_1	g_2	g_3
g_{-3}	1.00	0.35	0.24	0.10	0.08	0.08
g_{-2}	0.35	1.00	0.27	0.11	0.08	0.08
g_{-1}	0.24	0.27	1.00	0.15	0.15	0.13
g_1	0.10	0.11	0.15	1.00	0.33	0.30
g_2	0.08	0.08	0.15	0.33	1.00	0.41
g_3	0.08	0.08	0.13	0.30	0.41	1.00

$H(g_2) = 0.77$, $H(g_3) = 0.79$ for gaps in offers. We have not attempted to compute error bars for these values, but we would be very surprised if these differences are statistically significant. All these values demonstrate that the gap size is described by a long memory process, i.e. a stochastic process without a typical time scale. The fact that the granularity of the limit order book is well described by a long memory process indicates that the lack of liquidity in the market is not a temporary state, but rather one that persists in time.

The slow dynamics of the book as described by the persistence of the sparseness of the book and thus of the lack of liquidity suggests that when the first gap is large also the other gaps (on the same or on different side) are large. In other words the sparseness of the book is a property shared by the whole book *at the same instant of time.* To test this idea quantitatively we have computed the cross-correlation coefficient of the size of the first three gaps on both sides of the book. The result of this analysis is the symmetric correlation matrix shown in Table 1. First of all we note that all the correlation coefficients are positive, indicating a synchronous behavior of all the gaps. The correlation coefficient is significantly larger (of the order of $0.24 - 0.41$) when it is computed between gaps on the same side of the book. This is reasonable because it indicates that the sparseness of the book needs not to be shared simultaneously by both sides of the book. In fact at a given time the market can experience a liquidity crisis on the buy side, for example, but not on the sell side. Moreover it is worth noting that the correlation coefficient is stronger between the second and the third gap on both sides of the book. The data suggest that there may be a statistically significant correlation between gaps on different sides of the book. The 3σ level for the investigated sample is of the order of 0.004. However, these are long-memory processes, implying much larger error bars. For a long-memory process the error scale as $N^{-(1-H)}$. For a Hurst exponent of 0.75, for example, the errors scale as $N^{-0.25}$. Thus the standard error of 0.004 becomes roughly 0.065. Thus if these values are significant they are only just barely so. If real, this correlation indicates that a lack of liquidity on one side of the book makes a lack of liquidity on the other side more likely.

In conclusion we have shown that the multidimensional stochastic process which describes the state of the limit order book of a stock traded in a financial market has remarkable statistical properties. We have shown that the gaps closer to the best available prices are characterized by large fluctuations which are important

to explain large price fluctuations in terms of liquidity fluctuations. We have also shown that the liquidity availability described by the gaps size is a long memory process, indicating that the book dynamics lacks a typical time scale. Finally we have investigated the synchronous behavior of the size of different gaps showing that significant cross correlations between different gaps are present. These findings have relevant practical importance because they give some understanding of what determines financial risks, and gives some clues about how to reduce them.

Acknowledgments

References

[1] B. B. Mandelbrot, *The variation of certain speculative prices*, *J. Business* **36** (1963) 394–419.

[2] R. N. Mantegna and H. E. Stanley, *Scaling behavior in the dynamics of an economic index Nature* **376** (1995) 46–49.

[3] R. N. Mantegna and H. E. Stanley, *Introduction to econophysics: correlations and complexity in finance*, Cambridge University Press, Cambridge (1999).

[4] P. Gopikrishnan, M. Meyer, L. A. N. Amaral and H. E. Stanley, *Inverse cubic law for the distribution of stock price variations*, *Eur. Phys. J. B* **3** (1998) 139–140.

[5] P. K. Clark, *A subordinated stochastic process model with finite variance for speculative prices*, *Econometrica* **41** (1973) 135–256.

[6] X. Gabaix, P. Gopikrishnan, V. Plerou and H. E. Stanley, *A theory for power-law distributions in financial market fluctuations*, *Nature* **423** (2003) 267–270.

[7] J. D. Farmer, L. Gillemot, F. Lillo, S. Mike and A. Sen, *What really causes large price changes?*, *Quantitative Finance* **4** (2004) 383–397. Preprint available at xxx.lanl.gov/cond-mat/0312703.

[8] F. Lillo, J. D. Farmer and R. N. Mantegna, *Master curve for price-impact function*, *Nature* **421** (2003) 129–130.

[9] M. Potters and J.-P. Bouchaud, *More statistical properties of order books and price impact*, *Physica A* **324** (2002) 133–140.

[10] B. M. Hill, *A simple general approach to inference about the tail of a distribution*, *Annals Statistics* **3** (1975) 1163–1173.

[11] J. Beran, *Statistics for Long-Memory Processes*, Chapman & Hall (1994).

[12] J.-P. Bouchaud, Y. Gefen, M. Potters and M. Wyart, *Fluctuations and response in financial markets: the subtle nature of 'random' price changes*, *Quantitative Finance* **4** (2004) 176–190.

[13] F. Lillo and J. D. Farmer, *The long memory of efficient market*, *Studies in Nonlinear Dynamics and Econometrics* **8** (2004) 1.

[14] Z. Ding, C. W. J. Granger and R. Engle, *A long-memory property of stock market returns and a new model*, *J. Empirical Finance* **1** (1993) 83–106.

[15] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley and A. L. Goldberger, *Mosaic organization of DNA nucleotides*, *Physical Review E* **49** (1994) 1685–1689.