What really causes large price changes?

J. Doyne Farmer[1,5]**, László Gillemot**[1,2]**, Fabrizio Lillo**[1,3]**, Szabolcs Mike**[1,2] **and Anindya Sen**[1,4]

[1] Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA
[2] Budapest University of Technology and Economics, Budafoki út 8, H-1111 Budapest, Hungary
[3] Istituto Nazionale per la Fisica della Materia, Palermo University, Italy
[4] Mathematics Department, University of Chicago, Chicago, IL, USA

E-mail: jdf@santafe.edu, laci@santafe.edu, lillo@santafe.edu, mike@santafe.edu and anisen@santafe.edu

Abstract

We study the cause of large fluctuations in prices on the London Stock Exchange. This is done at the microscopic level of individual events, where an event is the placement or cancellation of an order to buy or sell. We show that price fluctuations caused by individual market orders are essentially independent of the volume of orders. Instead, large price fluctuations are driven by liquidity fluctuations, variations in the market's ability to absorb new orders. Even for the most liquid stocks there can be substantial gaps in the order book, corresponding to a block of adjacent price levels containing no quotes. When such a gap exists next to the best price, a new order can remove the best quote, triggering a large midpoint price change. Thus, 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. We show that this also explains price fluctuations on longer timescales. In addition, we present results suggesting that the risk profile varies from stock to stock, and is not universal: lightly traded stocks tend to have more extreme risks.

1. Introduction

It has been known for more than 40 years that price changes are fat-tailed [1, 15, 21, 36, 39–41, 43, 46–48, 50, 55], i.e. there is a higher probability of extreme events than in a normal distribution. This is important for financial risk, since it means that large price fluctuations are much more common than one might expect. There has been much conjecture concerning the origin of fat tails in prices [2, 3, 7, 8, 10, 20, 22, 30, 42, 44, 45, 60, 61]. Most of these theories are either generic mechanisms for generating power laws, such as multiplicative noise or maximization of alternative entropies, or agent-based

[5] Author to whom any correspondence should be addressed.

models that make qualitative predictions that are not very specific.

Two theories that deserve special mention because they make testable hypotheses about the detailed underlying mechanism are the subordinated random process theory advocated by Mandelbrot and Taylor [45] and also by Clark [8] and the recent theory of Gabaix *et al* [20]. Clark's proposal is that because order arrival rates are highly intermittent, aggregating in a fixed time interval leads to fat tails in price returns. Gabaix *et al*'s proposal is that high volume orders cause large price movements. We show that neither of these theories describes large price changes in the London Stock Exchange.[6]

Instead, we show that large price fluctuations are driven by fluctuations in liquidity, i.e. variations in the response of prices to changes in supply and demand. The number of agents that participate in the market at any given time, and thus the number of orders to buy or sell, is rather small. Even for a heavily traded stock, the typical number of orders on one side of the book at any given time is generally around 30. While it is in some cases a good approximation to regard the market as a statistical system, which can be treated using mathematical methods from statistical mechanics [5, 6, 13, 18, 57, 59], markets are far from the thermodynamic limit, and display strong finite size effects. Fluctuations in orders are important, but it is not the size of orders that drives large price changes, but rather the uniformity of their coverage of price levels. Revealed supply and demand curves at any instant in time are irregular step-like functions with long flat regions and large jumps. The market can be regarded as a granular medium, in which the incremental changes in supply and demand are the grains. The statistical properties of prices depend more on the fluctuations in revealed supply and demand than on their mean behaviour.

A common assumption is that large price returns $r(t)$ are asymptotically distributed as a power law. In more technical terms, letting $m(t)$ be the mid-price at time t, this means that $r(t) = \log m(t) - \log m(t - \tau)$ satisfies $P(r > x) \sim x^{-\alpha}$. $P(r > x)$ is the probability that $r > x$, τ is an arbitrary time interval, and $f(x) \sim g(x)$ means that $f(x)$ and $g(x)$ scale the same way[7] in the limit $x \to \infty$. α is called the *tail exponent*. It was initially thought that $\alpha < 2$, which is significant because this would imply that the standard deviation of price returns does not exist, and that, under aggregation, independent price returns should converge to a Levy stable distribution [15, 43]. However, most subsequent studies indicate that $\alpha > 2$ is more common [1, 36, 39–41, 46–48, 50, 55]. Nonetheless,

it still remains controversial whether a power law is always the best description of price returns. In this paper we do not attempt to resolve this debate. However, we will use power laws as a useful way to describe the asymptotic behaviour of price changes, in particular to compare the distribution of price changes for different stocks. Our results suggest that tail exponents vary from stock to stock, and are not universal.

In the remainder of this section we review the continuous double auction, explaining what we define as an 'event', and introducing notation that we will use throughout the rest of the paper. We also discuss the relationship between returns and events, and review related literature. Section 2 presents some summary statistics for the data set. Section 3 demonstrates that there is very little difference in the price response of large and small orders, and that large price fluctuations are driven by fluctuations in liquidity. This is made more explicit in section 4, where we show how gaps in the occupied price levels in the order book lead to large price changes, and that the gap distribution closely matches the return distribution. In section 5 we compare the returns triggered by market orders, limit orders, and cancellations, and show that they are very similar. Section 6 demonstrates that the tail behaviour of returns on the event scale matches the tail behaviour on longer time-scales, and that returns in event time and real time are similar. Section 7 studies the behaviour of the gap distribution in more detail, demonstrating that the (at least approximate) power law behaviour we observe is not driven by fluctuations in the number of orders in the book, but rather depends on correlations in occupied price levels. We conclude and offer a few speculations about the ultimate explanation of large price fluctuations in section 8.

1.1. Background: continuous double auction

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. (Note that the word 'quote' can be used either to refer to the limit price or to the limit order itself.) 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* or *asks*. 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

---

[6] For futher discussion of the paper of Gabaix *et al*, see [17].
[7] $f(x) \sim g(x)$ means that $\lim_{x \to \infty} L(x) f(x)/g(x) = 1$, where $L(x)$ is a slowly varying function. A slowly varying function $L(x)$ is a positive function that, for every t, satisfies $\lim_{x \to \infty} L(tx)/L(x) = 1$.

Figure 1. Schematic illustration of the continuous double auction mechanism. Limit orders are stored in the limit order book. We adopt the arbitrary convention that buy orders are negative and sell orders are positive. As a market order arrives, it has transactions with limit orders of the opposite sign, in order of price (first) and time of arrival (second). The best quotes at prices $a(t)$ or $b(t)$ move whenever an incoming market order has sufficient size to fully deplete the stored volume at $a(t)$ or $b(t)$, when new limit orders are placed inside the spread (when the limit price satisfies $b(t) < \pi < a(t)$) or when all the orders at the best prices are canceled. Bids and offers that fall inside the spread become the new best bids and offers.

the *inside quotes* or 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*, of size Δp. 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 figure 1. Because orders are placed for varying numbers of shares, matching is not necessarily one-to-one. For example, suppose the best offer is for 200 shares at \$60 and the next best is for 300 shares at \$60.25; a buy market order for 250 shares buys 200 shares at \$60 and 50 shares at \$60.25, moving the best offer $a(t)$ from \$60 to \$60.25. 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.

There are a variety of different order types defined in real markets, whose details differ from market to market. For our purposes here, any given order can always be decomposed into two types: we will call any component of an order that results in immediate execution an *effective market order*, and any component that is not executed immediately, and is stored in the limit order book, an *effective limit order*. For example, consider a limit order to buy whose limit price $\pi = a(t)$. Suppose the volume at $a(t)$ is 1000 shares, and the volume of the new limit order is 3000 shares. Then this limit order is equivalent to an effective market order for 1000 shares, followed by an effective limit order of 2000 shares with limit price $a(t)$.

In either case the same transactions take place, and the best prices move to $b(t + 1) = a(t)$ and $a(t + 1) = a(t) + g(t)$, where $g(t)$ is the price interval to the next highest occupied price level. Throughout the remainder of the paper we will simply call an effective limit order a 'limit order', and an effective market order a 'market order'.

When a market order arrives it can cause changes in the best prices. This is called *market impact* or *price impact*. Note that the price changes are always in the same direction: a buy market order will either leave the best ask the same or make it bigger, and a sell market order will either leave the best bid the same or make it smaller. The result is that buy market orders can increase the mid-price $m(t) = (a(t) + b(t))/2$, and sell orders can decrease it.

There is no unique notion of price in a real market. We will let π be the *limit price* of a limit order, and $m(t) = (a(t) + b(t))/2$ be the *midpoint price* or *mid-price* defined by the best quotes. All the results of this paper concern the mid-price, rather than transaction prices, but at longer timescales this makes very little difference, since the mid-point and transaction prices rarely differ by more than half the spread. The mid-point price is more convenient to study because it avoids problems associated with the tendency of transaction prices to bounce back and forth between the best bid and ask. Price changes are typically characterized as *returns* $r_\tau(t) = \log m(t) - \log m(t - \tau)$.

1.2. What causes returns?

In this paper we study changes in the mid-price at the level of individual events. The arrival of three kinds of events can cause the mid-price to change:

- *Market orders*. A market order bigger than the opposite best quote widens the spread by increasing the best ask if it is a buy order, or decreasing the best bid if it is a sell order.
- *Limit orders*. A limit order that falls inside the spread narrows it by increasing the best bid if it is a buy order, or decreasing the best ask if it is a sell order.
- *Cancellations*. A cancellation of the last limit order at the best price widens the spread by either increasing the best ask or decreasing the best bid.

We prefer to study individual events for several reasons. (1) It removes any ambiguity about time scale, and makes it easier to compare stocks with different activity levels. (2) It minimizes problems associated with clustered volatility (the positive autocorrelation of the absolute value of price changes). Clustered volatility is also driven by variations in event arrival rates, so its effect is weaker when the time lag between price changes is measured in terms of the number of intervening events. (3) Individual events are the most fundamental level

of description—returns on any longer scale can be constructed from returns on an event scale. We will generally measure time in terms of *event time*: the time interval between events t_1 and t_2 is measured as the number of intervening events plus one. (One is added so that two adjacent events comprise an event time interval of one.)

Analysis at the event scale is particularly useful under the assumption that large price fluctuations are asymptotically power law distributed because of the invariance properties of power laws under aggregation. If two independent variables X and Y both have power law tails with exponents α_X and α_Y, then the variable $Z = X + Y$ also has a power law tail with exponent $\min(\alpha_X, \alpha_Y)$. Thus, if we show that returns at the level of individual events have power law tails with exponent α and if there are no strong and persistent correlations, returns on any longer event time scale should also have power law tails with exponent α. In fact, α appears to be the same whether returns are measured in event time or real time (see section 6).

1.3. Review of previous work

There is considerable prior work using limit order data to address questions about market microstructure. For example, in a very early study, Niederhoffer and Osborne discussed the granularity of revealed supply and demand, and showed orders tend to cluster at particular prices [49]. Clustering of limit orders, as well as stop-loss and take-profit orders, is also studied in [32, 33, 51]. Of particular relevance is the work of Biais *et al* [4], who, for the Paris Bourse, document several properties of order flow, including the concentration of orders near the best prices. They study the relation between order flow and the dynamics of prices, and mention the existence of gaps in the limit order book. Another relevant observation was made by Knez and Ready [35] and Petersen and Umlauf [53], who demonstrate that, for the New York Stock Exchange (NYSE), the most important conditioning variable for price impact is the size of an order relative to the volume at the best price. Goldstein and Kavajecz [23] document the effect of changes in tick size on limit order book depth, while several theoretical papers justify the existence of a positive bid–ask spread and investigate the motivations for placing limit orders as opposed to market orders [9, 19, 25, 29, 31, 52]. In addition, we should mention recent empirical studies by Coppejans and Domowitz [11] and Coppejans *et al* [12] that document variations in liquidity and co-movements of liquidity with returns and volatility.

Our work adds to this literature by investigating the relationship between order placement and price movement at the level of individual events. Our motivation is to understand what drives large price movements, to gain insight into the fat tails of price returns. A relevant paper

Table 1. Summary statistics of the 16 stocks we study for the period 1999–2002. The columns give the number of events of each type, in thousands. All events are 'effective' events—see the discussion in section 1.1

Tick	Market orders	Limit orders	Cancellations	Total
AZN	652	2067	1454	4173
BAA	226	683	487	1397
BLT	297	825	557	1679
BOOT	246	711	501	1458
BSY	404	1120	726	2250
DGE	527	1329	854	2709
GUS	244	734	518	1496
HG.	228	676	472	1377
LLOY	723	1664	1020	3407
PRU	448	1227	821	2496
PSON	373	1063	734	2170
RIO	381	1122	771	2274
RTO	276	620	389	1285
RTR	479	1250	820	2549
SBRY	284	805	561	1650
SHEL	717	4137	3511	8365
Total	6505	20 033	14 196	40 734

in this regard is the work of Plerou *et al* [54], who showed that, for the NYSE in a fixed time interval, the scaling behaviour in the standard deviation of individual price fluctuations roughly matches that of price fluctuations, and dominates over fluctuations in the number of events. This seems to contradict the later conclusions of the same authors in Gabaix *et al* [20].

2. Data

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 analysed 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) and Shell Transport & Trading Co. (SHEL). These stocks were selected because they have high volume and they are all continuously traded during the full period. Table 1 gives a summary of the number of different events for the 16 stocks.[8]

The London Stock Exchange consists of two parts: the completely automated electronic downstairs market (SETS) and the upstairs market (SEAQ). The trading volume is split roughly equally between the two markets. We study the downstairs market because we have a record of each action by each trader as it occurs. In contrast,

---

[8] We have observed similar results on less liquid stocks, except that the statistics tend to be poorer, and as we discuss in section 4, e.g. figure 8, less liquid stocks appear to have lower tail exponents.

trades in the upstairs market are arranged informally between agents, and are printed later. There are no designated market makers for SETS; however, any member of the exchange is free to act as a market maker by posting simultaneous bids and offers. This should be contrasted with the NYSE, for example, which has a designed specialist to trade each stock.[9] During the period we study the book is fully transparent, i.e. all orders in the book are fully revealed.[10]

Trading begins each day with an opening auction. There is a period leading up to the opening auction in which orders are placed but no transactions take place. The market is then cleared and for the remainder of the day (except for occasional exceptional periods) there is a continuous auction. We remove all data associated with the opening auction, and only analyse orders placed during the continuous auction.

An analysis of the limit order placement shows that, in our dataset, approximately 35% of the effective limit orders are placed *inside the book* ($\pi > a(t)$ or $\pi < b(t)$). Thirty-three percent are placed at the best prices ($\pi = b(t)$ or $\pi = a(t)$), and 32% are placed inside the spread ($b(t) < \pi < a(t)$). This is roughly true for all the stocks except for SHEL, for which the percentages are 71%, 18% and 11%, respectively.[11] Moreover, for all the stocks the properties of buy and sell limit orders are approximately the same.

In this dataset, cancellation occurs roughly 32% of the time at the best price and 68% of the time inside the book. This is quite consistent across stocks and between the cancellation of buy and sell limit orders. The only significant deviation is once again SHEL, for which the percentages are 14% and 86%.

3. Fluctuations in liquidity drive the tails

The assumption that large price changes are caused by large market orders is very natural. A very large market order will dig deeply into the limit order book, causing transactions at many price levels, increasing the spread, and changing the mid-price. Surprisingly, this is *not* the cause of most large price changes. Instead, as we will demonstrate in this section, most large price changes are due to discrete fluctuations in liquidity, manifested by gaps in filled price levels in the limit order book. Large price changes caused by large orders are very rare,

and play an insignificant role in determining the statistical properties of price changes.

In this section we will focus on price changes caused by market orders, and in section 5 we will discuss price changes due to limit orders and cancellations.

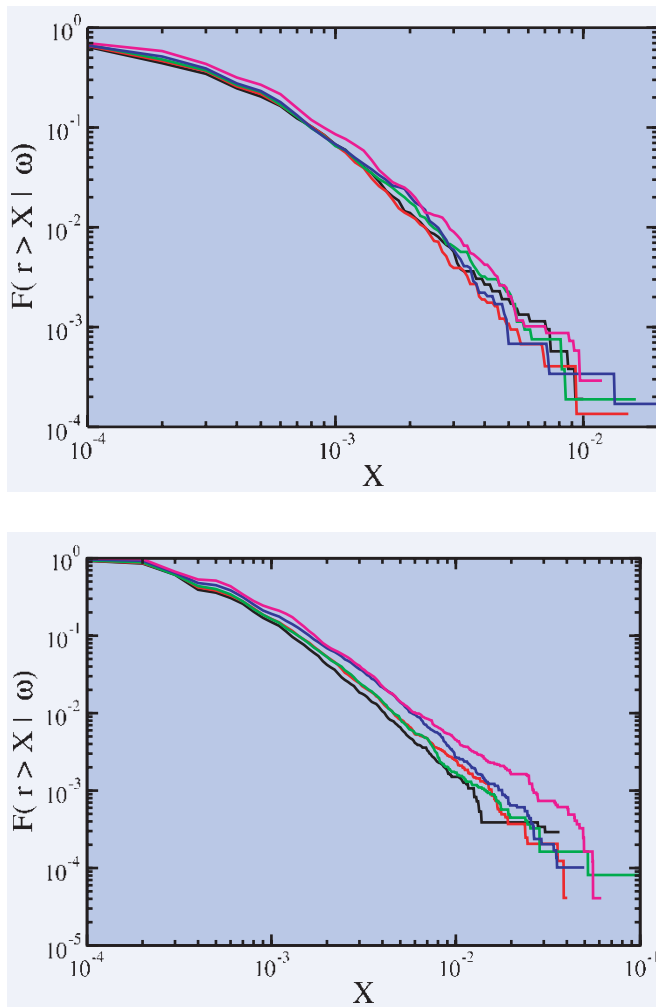## 3.1. Large returns are not caused by large orders

We first demonstrate that most large returns are not caused by the arrival of large market orders. The probability density function of price returns that are caused by market orders can trivially be written as $p(r) = \int p(r|\omega)p(\omega) d\omega$, where $p(r)$, $p(\omega)$ and $p(r|\omega)$ are the probability density functions for returns r, market order size ω, and returns given market order size. The conditional probability $p(r|\omega)$ characterizes the response of prices to new orders, and can be viewed as the probability density of market impacts, or, alternatively, as characterizing the distribution of liquidity for market orders. When a market order of size ω arrives the mid-price will move if ω is larger than or equal to the volume at the matching best price (i.e. the bid for sell market orders and the ask for buy market orders). In the limit of continuous prices we can trivially write

$$p(r|\omega) = (1 - g(\omega))\delta(r) + g(\omega)f(r|\omega), \qquad (1)$$

where $\delta(\cdot)$ is the Dirac delta function[12] and $g(\omega)$ is the probability that the mid-price moves as a function of the order size ω. The function $f(r|\omega)$ is the probability of a price shift r, conditioned on the mid-price moving in response to an order of volume ω. $g(\omega)$ and $f(r|\omega)$ behave quite differently. $g(\omega)$ depends strongly on ω. For the LSE it scales roughly as $g(\omega) \sim \omega^{0.3}$, about the same as the average market impact [17]. In contrast, $f(r|\omega)$ is surprisingly independent of the volume ω, in the sense that the unconditional fluctuations in r dominate the dependence on ω.

To demonstrate this, in figure 2 we show the cumulative probability for non-zero price returns conditioned on order size, i.e. $F(r > X|\omega) = \int_X^\infty f(r|\omega) dr$, for several different ranges of market order size ω. The orders are sorted by size into five groups with roughly the same number of orders in each group. The distributions for each range of ω are roughly similar, both for individual stocks such as AZN, and for the pool. Each curve approximately approaches a power law for large r independent of ω, illustrating that the key property determining large price returns is fluctuations in market impact, and that the role of the volume of the order initiating a price change is minor. For the pooled stock result, for large returns the curves for large order size tend to be on top of those for small order size, illustrating a weak dependence on order size, but this is relatively
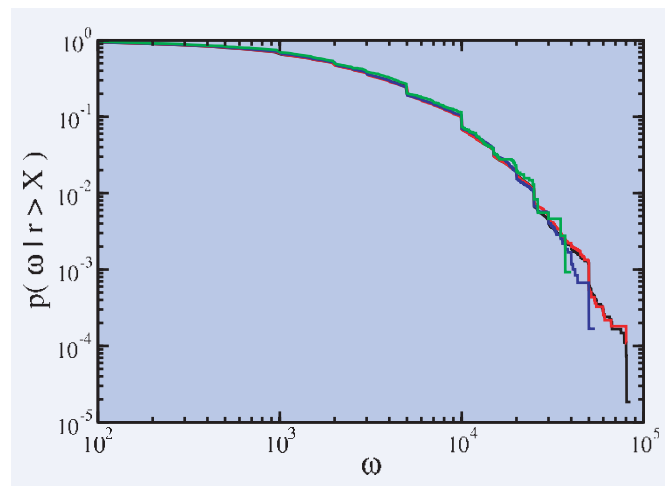
---

[9] Another difference between the two markets is that clearing in the LSE is fully automated and instantaneous; in contrast, in the NYSE, clearing is done manually, creating an uncertainty in response times.
[10] In 2003 the LSE began to allow 'iceberg orders', which contain a hidden component that is only revealed as the exposed part of the order is removed.
[11] For the Paris Stock Exchange from 1994, Bias *et al* [4] observed 42% of the orders inside the spread, 23% at the best, and 35% inside the book. We do not know why SHEL is anomalous, although it is worth noting that it is the most heavily traded stock in the sample, a close proxy is also traded on the NYSE, and it has the second largest tail exponent among the 16 stocks chosen for the study.

[12] The Dirac delta function $\delta(x)$ is defined so that $\delta(x) = 1$ when integrated over any domain that includes 0, and $\delta(x) = 0$ otherwise.

Figure 3. $p(\omega|r > X)$, the distribution of market order sizes conditioned on generating a return greater than X, for the stock AZN. The values of X correspond to the 50 percentile (red), 90 percentile (blue), and 99 percentile (green) of the return distribution. The black line is the unconditional distribution $p(\omega)$. The jumps are due to the tendency to place orders in round numbers of shares.

Figure 2. Dependence of returns on order size. $F(r > X|\omega)$ is the probability of a return $r > X$ conditioned on the order size ω and on the fact that the price shift is non-zero. Results shown are for buy orders, but similar results are seen for sell orders. The orders are sorted by size into five groups with roughly the same number of orders in each group. Ranging from small orders to large orders, the curves are black, red, green, blue, and magenta. In panel (a) we show the result for AZN and in panel (b) we show the average over the pool of 16 stocks described in table 1. For the pooled data for each stock we normalize the order volume to the sample mean, and then combine the data. Each curve approximately approaches a power law for large r independent of ω, illustrating that the key property determining large price returns is fluctuations in market impact, and that the role of the volume of the order initiating a price change is minor.

small, and not visible in the results for individual stocks. Although we do not present them here, using the TAQ database we have obtained similar results for a small sample of stocks traded in the NYSE, suggesting that this behaviour is not specific to the LSE.[13] This is particularly interesting given the significant differences in

the structure of the NYSE, and also because the TAQ data set includes upstairs trades.

To reinforce this point, in figure 3 we show $p(\omega|r > X)$, the probability of a market order of size ω conditioned on the return being greater than a certain threshold X. We do this for the stock Astrazeneca (AZN) for several different values of X, getting virtually the same curve independent of X. The distribution of order sizes that generate large returns is essentially the same as those that generate small returns. Similar results are obtained for all the stocks in our sample. Thus, it seems that order size does not play an important role in generating large returns.

Gabaix *et al* [20] have recently proposed that large returns can be explained by assuming the market impact density function $p(r|\omega)$ is sharply peaked around a central value[14] $k\omega^{1/2}$, so that it can be approximated as a Dirac delta function $\delta(r - k\omega^{1/2})$. The result of figure 2 makes it clear that the sharply peaked assumption is a poor approximation—the distribution is quite broad, in the sense that the conditional distribution $p(r|\omega) \approx p(r)$, and $p(r)$ is not sharply peaked.

[13] Note added in press: based on 5 min averages, Weber and Rosenow [64] also report that liquidity is the dominant effect, using data from NYSE and Island.

[14] Analyses of several markets make it clear that the mean response of prices to orders varies considerably from market to market, and is not in general well characterized by a square root law [17, 18, 38, 57]. It is also worth noting that the volume distribution in figure 3 does not appear to have a power law tail. This is true of all the stocks in our sample. In contrast, power law tails are observed for stocks in the NYSE [24], and this was also observed to be true for upstairs data in the LSE (Lillo and Farmer, unpublished). See also the discussion in [37].

Table 2 . Summary table of the percentage of the time that non-zero changes in the best prices are equal to the first gap (left) and that the market order volume ω exactly matches the volume at the corresponding best price V_{best}(right). Assuming a Bernoulli process the sample errors are of the order of 0.1–0.2%.

	% of non-zero returns equal to first gap		% of non-zero returns with $\omega = V_{\text{best}}$	
Tick	Sell	Buy	Sell	Buy
AZN	94.3	99.6	83.7	90.1
BAA	95.9	99.1	86.2	87.0
BLT	95.0	99.2	85.8	85.6
BOOT	95.9	99.2	85.7	84.7
BSY	94.3	99.7	85.0	88.0
DGE	96.3	99.7	86.5	87.4
GUS	95.8	99.5	85.0	83.5
HG.	95.9	99.5	85.8	83.0
LLOY	97.3	99.8	88.4	88.6
PRU	95.9	99.5	85.5	78.1
PSON	93.1	99.6	81.2	86.2
RIO	95.7	99.7	84.6	86.4
RTO	96.1	99.5	84.4	85.3
RTR	93.0	99.7	83.5	85.4
SBRY	95.6	99.4	85.3	83.2
SHEL	98.7	99.9	93.0	92.8
Average	95.6	99.5	85.6	86.0

3.2. Volume dependence of mean market impact

From a variety of previous studies it is clear that the mean market impact is an increasing function of ω [14, 16, 26, 27, 34, 38, 56–58, 62, 63]. How do we reconcile this with our claim that the distribution of market impact is almost independent of volume? The key is that the main effect of changing the order volume is to change the probability that the price will move, with very little effect on how much it moves. From equation (1) one easily obtains

$$E(r|\omega) = g(\omega) \int f(r|\omega) r \, dr \propto g(\omega), \qquad (2)$$

where for the last proportionality relationship we have used the result from figure 2 that $f(r|\omega)$ is almost independent of ω. Thus, the expected price change scales as the probability of a price change, a relationship that we have verified for both the LSE and the NYSE. However, this variation is still small in comparison with the intrinsic variation of returns; the mean market impact of a very large order is less than the average size of the spread, but the largest market impacts are often more than 10 times this large.

3.3. Correlations between order size and liquidity

One possible explanation for the independence of price response and order size is that there is a strong correlation between order size and liquidity. There is an obvious strategic reason for this: agents who are trying to transact large amounts split their orders and execute them a little at a time, watching the book, and taking whatever liquidity is available as it enters. Thus, when there is a lot of volume in the book they submit large orders, and when there is less volume, they submit small orders. This effect tends to even out the price response of large and small orders. We will see that this effect indeed exists, but it is only part of the story, and is not the primary determinant of the behaviour we observe here.

In fact, the unconditional correlation between market order size and volume at the best is rather small. For AZN, for example, it is about 1%. However, if we restrict the sample to orders that change the mid-price, the correlation soars to 86%. The reason for this is that for orders that do not change the price, there is essentially no correlation between order size and volume at the best. For the rarer case of orders that do change the price, in contrast, most market orders exactly match the volume at the best. As shown in table 2, for the stocks in our sample, 86% of the buy orders that change the price exactly match the volume at the best price. (This is 85% for sell orders.)

The relationship between the volume of market orders and the best price becomes more evident with a nonlinear analysis. Figure 4 shows $E(\omega|V_{\text{best}})$, where ω is the market order size and V_{best} is the volume at the corresponding best price. We see in this figure that the expected order size is non-zero even for the smallest values of V_{best}. It grows monotonically with V_{best}, but with a slope that is substantially less than one, and a roughly concave shape. This makes the nonlinear correlation between order size and liquidity clear. However, in the following section we will see that the dependence of order size on liquidity is not strong enough to substantially suppress large price fluctuations.
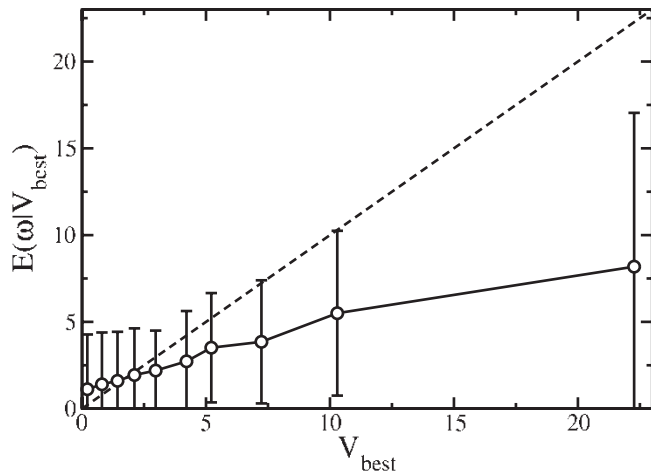
Figure 4. The dependence of order size on liquidity. Liquidity is measured as the volume at the best price, which is binned into deciles with roughly equal numbers of events. The vertical axis shows the mean market order volume for each decile, and the horizontal shows the volume at the best. The units of both axes are thousands of shares. The ranges shown around each point indicate one standard deviation (and are not standard errors). The dashed line has slope 1 and serves as a point of comparison.

3.4. Liquidity fluctuations drive price fluctuations

In this section we demonstrate in concrete terms that price fluctuations are driven almost entirely by liquidity fluctuations. To do this we study the virtual market impact, which is a useful tool for probing the supply and demand curves defined by the limit order book. Whereas the true market impact $p(r|\omega)$ tells us about the distribution of impacts of actual market orders, as discussed in section 3.1, the virtual market impact is the price change that *would* occur at any given time if a market order of a given size were to be submitted. More formally, at any given time t the limit orders stored in the order book define a revealed supply function $S(\pi, t)$ and revealed demand function $D(\pi, t)$. Let $V(\pi, t)$ be the total volume of orders stored at price π. The revealed supply function is

$$S(\pi, t) = \sum_{i=a(t)}^{\pi} V(i, t). \tag{3}$$

The revealed supply function is non-decreasing, and so for any fixed t has a well-defined inverse $\pi(S, t)$. The virtual market impact is the price shift caused by a hypothetical order of size S, e.g. for buy orders it is $\pi(S, t) - a(t)$. The virtual market impact for sell market orders can be defined in terms of the revealed demand in a similar manner. By sampling at different values of t, for any fixed hypothetical order size we can create a sample distribution of virtual market impacts. This naturally depends on the sampling times, but these can be chosen to match any given set of price returns.

In figure 5 we show the cumulative distribution of virtual market impacts for the stock AZN for several
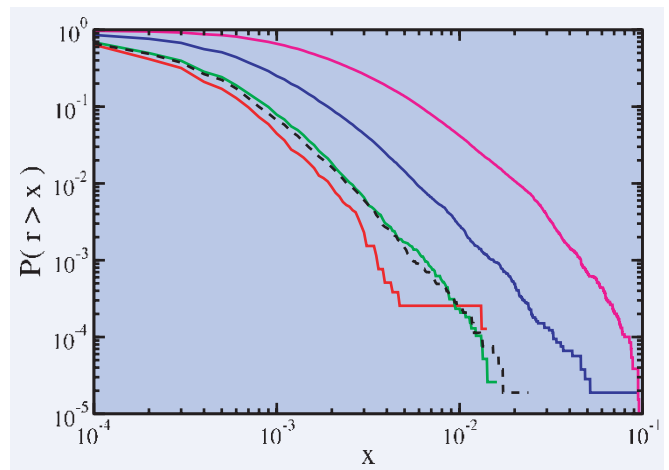
Figure 5. The cumulative probability distribution $P(r > x)$ of virtual market impacts for the stock AZN, based on samples of the limit order book taken just before the arrival of a market order. From left to right the continuous lines are the virtual market impact for hypothetical buy orders of size $\omega = 100$ shares (red), $\omega = 1600$ shares (green), $\omega = 8700$ shares (blue), and $\omega = 25\,000$ shares (magenta); these correspond to the 0.1, 0.5, 0.9, and 0.99 quantiles of market order size. These are compared to the average distribution of returns (black dashed), which is very similar to the virtual market impact for the 0.5 quantile. Note that we have discarded cases in which the virtual market impact is undefined due to excessively large hypothetical market order volume, but these comprise only about 3% of the events.

different values of D, corresponding to different quantiles of market order size. The cumulative distributions define a set of approximately parallel curves. They are shifted as one would expect from the fact that larger hypothetical order sizes tend to have larger virtual market impacts. These curves are similar in shape to the distribution of returns. Most striking, for the median market order size, the curves are almost identical. We see similar results for all the stocks in our sample. This demonstrates quite explicitly that the distribution of returns is determined by properties of the limit order book, and that the typical price return corresponds to the price response to an order of typical size. The fact that the price distribution is so close to the virtual market impact of a typical order shows that correlations between order size and liquidity are not important in determining price fluctuations.

4. Granularity of supply and demand

At first sight, the behaviour described in the previous section seems baffling: how can market order size be so unimportant to price response? In this section we show how this is due to the granularity of revealed supply and demand, which causes large fluctuations in liquidity.

The cause of this puzzling behaviour is fluctuations in occupied price levels in the limit order book. In particular, one can define the size of the *first gap g* as the absolute difference between the best log price, π_{best},
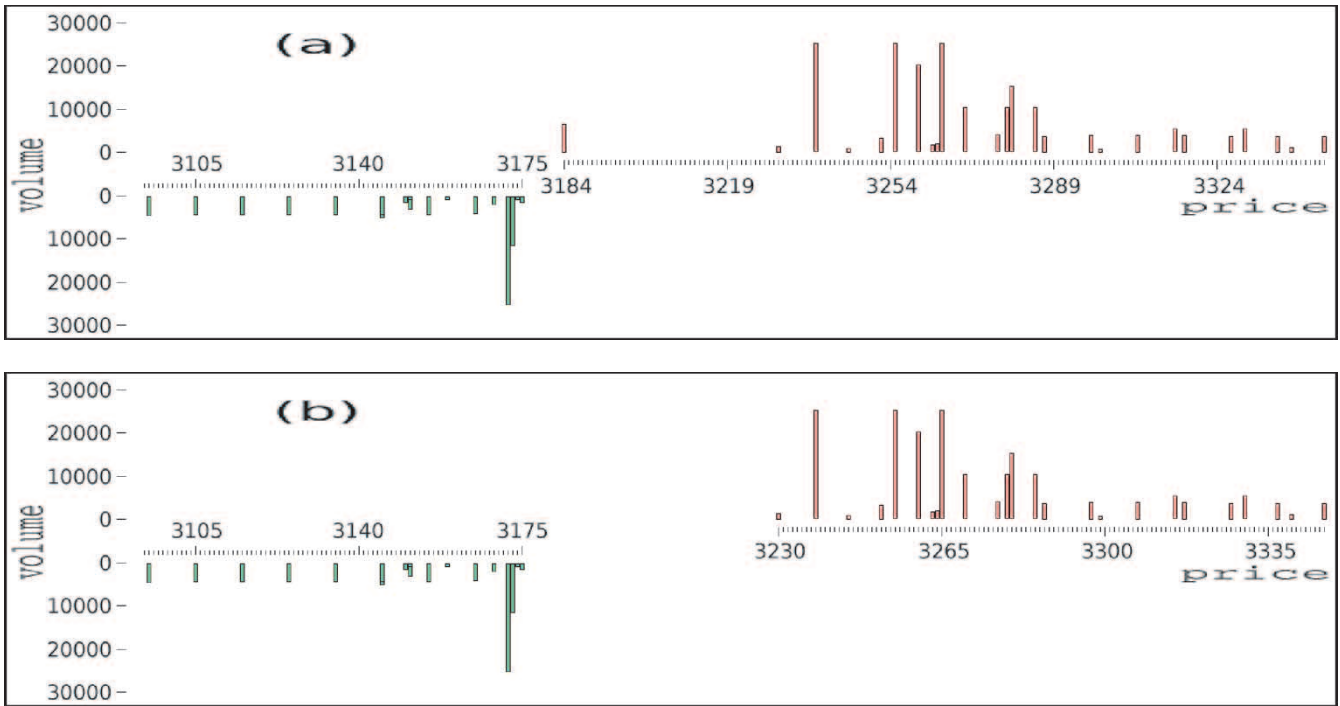
Figure 6. A typical configuration of the limit order book for AZN before and after a large price fluctuation. The two panels plot the volume (in shares) of limit orders at each price level; sell limit orders are shown as positive, and buy limit orders as negative. In panel (a) we see that there is a large gap between the best ask price and the next highest occupied price. The arrival of a market order to buy removes all the volume at the best ask, giving the new limit order book configuration shown in panel (b), which has a much higher best ask price than previously.

and the log price of the next best quote, π_{next}, as $g = |\log \pi_{\text{best}} - \log \pi_{\text{next}}|$. In figure 6 we show a typical set of events before and after a large price fluctuation. In figure 6(a) we see that there is a large first gap on the sell side of the limit order book. In figure 6(b) we see the configuration of the book an event later, after a market order of exactly the same size has removed all the volume at the best ask. This results in a large change in the mid-point price. Thus, the large return simply reflects a large gap inside the book that has been revealed by the removal of the best ask. We find that this is the typical behaviour underlying almost all large returns.

As already shown in table 2, typically about 85% of market orders that result in price changes exactly match the volume at the best price. Furthermore, as shown in the first two columns of this table, when they do exceed the volume at the best price, it is quite rare that they penetrate the next occupied level. For buy (sell) market orders, on average 95.5% (99.5%) of the non-zero shifts in the best price are exactly equal to the first gap. Interestingly, there is a significant asymmetry between buy and sell. Traders act to minimize their transaction costs, so that jumps of more than one gap are rare, particularly for sell orders. (However, as we stressed earlier, the return distribution can be generated by a constant order of median size—this correlation is interesting, but not essential.) Note that the trader initiating the change *does not* pay a large spread—that

would only happen to the next trader, if she were to immediately place a market order.[15]

This is by far the most common scenario that generates large price changes. In figure 7(a) we compare the distribution of the first gaps to the distribution of price returns for the stock AZN. We see that the distributions are very similar. For the first gap size the tail index $\alpha = 2.52 \pm 0.07$, and for the return distribution $\alpha = 2.57 \pm 0.08$, showing that the scaling behaviours are similar. However, the similarity is not just evident in the scaling behaviour—the match is good throughout the entire range, illustrating that most of the large price changes are caused by events of this type. Figure 7(b) shows the same comparison for the pool of 16 stocks. The agreement in this case is even more striking.[16]

To demonstrate that the correspondence in the above figure is not just a coincidence for AZN, we have computed the tail exponents for returns and first gap size for all 16 stocks in our data set. We do this using

---

[15] In general, after a large shift in the bid or ask price, the next orders tend to be limit orders, but we have not yet been able to study the statistical properties of the sequence of subsequent events in detail. See [4, 9, 25, 29, 52] for an empirical study of the role of quote size after a market order.
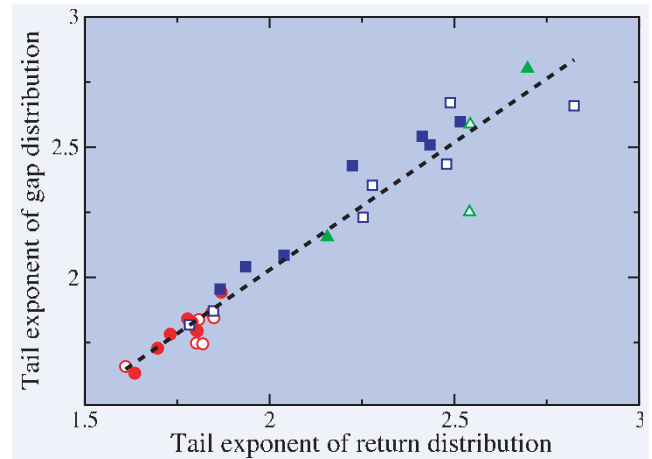
[16] We have also studied the distribution of higher order gaps. Moving away from the best price π_0, the nth order gap for $n = 1, 2, \ldots$ can be recursively defined as $g_n = |\log \pi_{n-1} - \log \pi_n|$, where π_n is the nth occupied price level. Interestingly, we find that the tail behaviour of higher order gaps is the same as that of g_1.
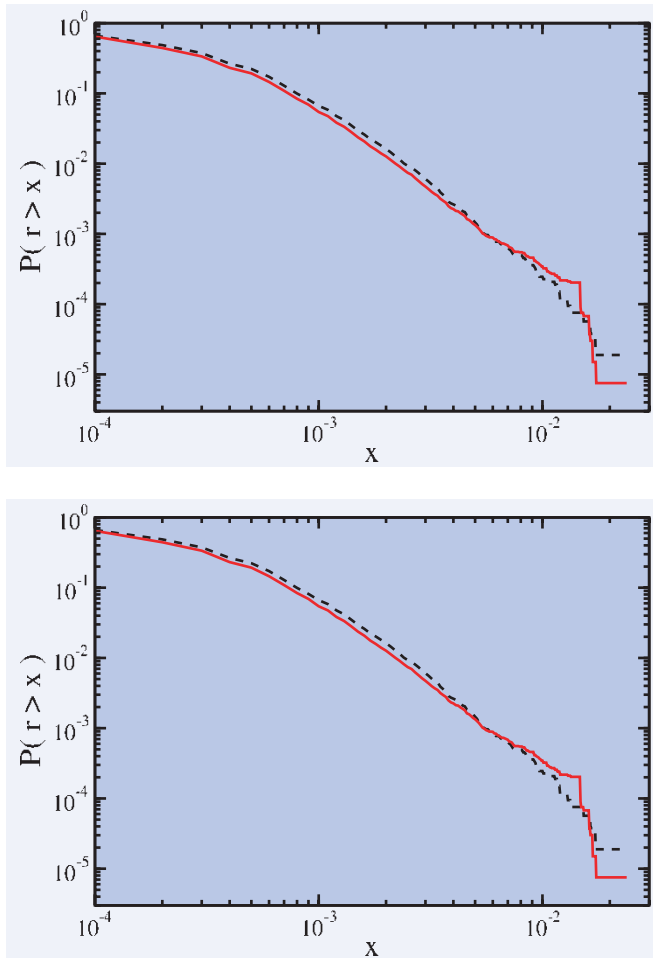
Figure 7. The cumulative distribution $P(g > x)$ of the size of first gaps g (red continuous line), compared to the cumulative distribution of returns generated by market orders $P(r > x)$ (black dashed line). Panel (a) refers to buy market orders for AZN, on a double logarithmic scale to highlight the tail behaviour. The two distributions are very similar. The result is even more impressive when we consider the average over the 16 stocks described in table 1 (Panel (b)).

Figure 8. A comparison of tail exponents α for returns (horizontal axis) vs. tail exponents for the first gap (vertical axis). The first gaps are sampled immediately preceding the market order that triggers the return. This is done for all the stocks in table 1, and includes both the buy and sell sides of the limit order book. We use red circles for low liquidity stocks, blue squares for medium liquidity stocks, and green triangles for high liquidity stocks. Empty symbols refer to sell market orders and filled symbols to buy market orders. The black dashed line is the linear regression. We see that there is a clear positive relationship between the number of orders and the tail exponent.
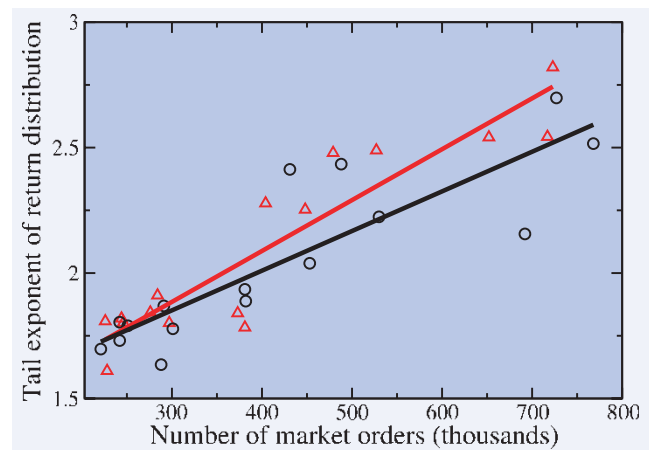
Figure 9. Dependence of the tail exponent of price changes on the number of market orders, for the data of figure 8. Black circles are for positive returns (caused by buy market orders), and red triangles are for negative returns (caused by sell market orders). The steeper red curve is the fit to negative returns and the black curve the fit to positive returns.

a Hill estimator [28] by considering the largest $\sqrt{n}$ returns, where n is the size of the sample. The results are shown in figure 8, where we plot the tail exponents of returns against those of first gaps. The points cluster tightly along the diagonal, making it clear that there is a strong positive correlation ($R^2 = 0.93$) between the gap tail exponents and the return tail exponents; a least-squares linear fit has a slope of 0.98 ± 0.05, and we are unable to reject the null hypothesis that the tail exponents of the gaps and returns are drawn from the same distribution.

It is worth noting that lightly traded stocks tend to have a smaller tail exponent than heavily traded stocks. This is already clear in figure 8, where we have used a colour coding to identify stocks of different volume. To investigate this more quantitatively, in figure 9 we plot the tail exponent of price changes against the number of market orders for each stock in the sample. While the relationship is noisy, there is a clear positive trend; the

slopes of the linear fits to positive and negative returns are both highly statistically significant. The tail exponents vary from about 1.6 to 2.8, whereas the error bars are more than a factor of 10 smaller than the range of variation.[17] This suggests that the tail exponent is not universal. Not surprisingly, stocks that are more heavily

---

[17] Table 3 gives error bars for tail exponent estimations based on absolute returns (including both buy and sell events), while figures 7 and 8 are based on buy and sell returns taken separately, but the error bars are comparable.

traded tend to display less extreme risk than stocks that are lightly traded.

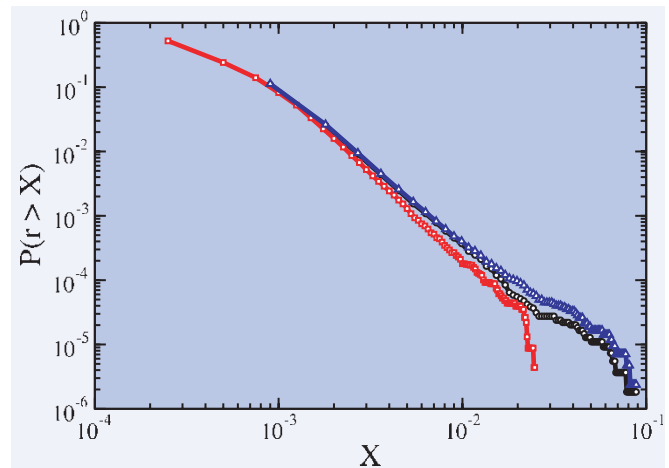## 5. Limit orders and cancellations

We have so far considered only returns caused by market orders. In this section we will discuss returns caused by limit orders and cancellations, and show that they are statistically indistinguishable from those caused by market orders.

The distribution of returns caused by limit orders, market orders, and cancellations for AZN is shown in figure 10. The distributions are quite similar. To make the analysis more quantitative, we compute the tail exponents for the 16 stocks in the sample. Table 3 shows the Hill estimates for returns caused by market orders, limit orders and cancellations. The table shows clearly that the tail exponents for the three distributions are very close: For all 16 stocks the estimated tail exponents are well within the 95% confidence intervals. On the other hand, table 3 shows that there is a much larger variation across stocks.

Why are the distributions of different events so similar? The correspondence between returns caused by market orders and returns caused by cancellations is not surprising. In terms of the effect on best prices, removal of the volume at the best price by cancellation is equivalent to removal by a market order, in both cases creating a price change equal in size to the first gap. However, for a limit order this is not so obvious: a limit order that falls inside the spread decreases the spread, and causes a price change in the opposite direction from market orders and cancellations. To investigate this we have studied the distribution of the spread, which also appears to be a power law, but with a larger tail exponent. At this point the reason for the close correspondence between the returns generated by limit and market orders remains unexplained.

6. Price changes on longer timescales

One can naturally ask whether the microscopic event level analysis we have presented here explains the statistical properties of price changes over longer and/or fixed time horizons. The timescale for the analyses we have presented here is quite fast: for Astrazeneca, for example, within the period of this dataset there are on average about eight market orders every 5 min, but due to highly uneven rates of event arrival, it is not uncommon that many events arrive within the same second. It is natural to question whether events on this timescale reflect the properties of longer timescales, e.g. on the daily timescale of many other studies. Given that order arrival is highly clustered in time, it is also natural to ask whether a description in event time also provides an explanation in real time. We now address both of these questions.


Figure 10. A comparison of the distribution of absolute returns (including both buy and sell orders) caused by limit orders, market orders, and cancellations for AZN. The black circles are for market orders, the blue triangles are for limit orders, and the red squares are for cancellations.

Table 3. Hill estimator of the return caused by market orders, limit orders and cancellations. We analysed the top percentile and give 95% confidence intervals. The confidence intervals assume the samples are uncorrelated, and so may be over-optimistic.

Tick	Market orders	Limit orders	Cancellations
AZN	2.57 ± 0.08	2.56 ± 0.07	2.57 ± 0.06
BAA	1.77 ± 0.06	1.77 ± 0.08	1.76 ± 0.08
BLT	1.8 ± 0.1	1.8 ± 0.1	1.8 ± 0.1
BOOT	1.80 ± 0.07	1.80 ± 0.07	1.79 ± 0.08
BSY	2.28 ± 0.04	2.25 ± 0.06	2.26 ± 0.06
DGE	2.31 ± 0.06	2.31 ± 0.07	2.31 ± 0.07
GUS	1.80 ± 0.09	1.80 ± 0.07	1.80 ± 0.08
HG.	1.7 ± 0.1	1.7 ± 0.1	1.66 ± 0.07
LLOY	2.72 ± 0.03	2.72 ± 0.04	2.72 ± 0.05
PRU	2.20 ± 0.06	2.20 ± 0.07	2.20 ± 0.06
PSON	1.9 ± 0.1	1.9 ± 0.1	1.91 ± 0.09
RIO	1.83 ± 0.09	1.82 ± 0.08	1.83 ± 0.07
RTO	1.73 ± 0.09	1.73 ± 0.08	1.73 ± 0.07
RTR	2.45 ± 0.04	2.44 ± 0.05	2.45 ± 0.03
SBRY	1.89 ± 0.07	1.88 ± 0.07	1.88 ± 0.07
SHEL	2.62 ± 0.04	2.61 ± 0.07	2.62 ± 0.04

The permanence of these price movements can be seen through the continuity between price returns on the single event scale and the multiple event scale. Figure 11 shows the return density function for AZN for 1, 2, 4, 8, 16, and 32 market order arrivals, as well as for 5, 10, and 20 min timescales. Each curve consists of a non-power law central behaviour crossing over to an approximate power law tail behaviour. We see that the tails of all the distributions are quite similar, though the crossover point from the central behaviour to the tail behaviour increases for longer timescales. The regularity of the movement in the crossover point as we vary the timescale, and the similarity of the tail behaviour across a range of timescales, indicates that the tail properties on the single
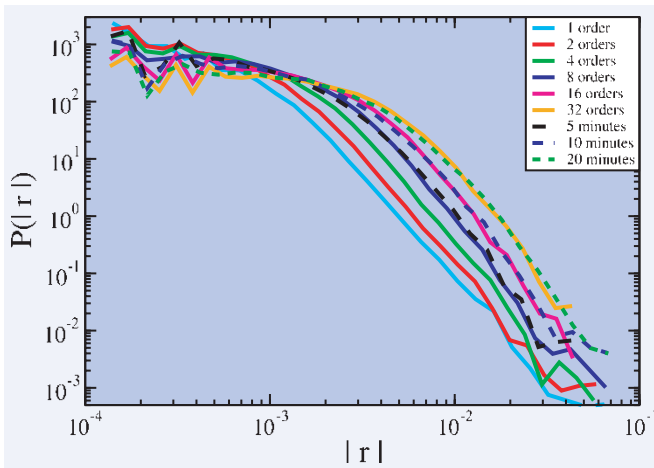
Figure 11. Price returns aggregated on different timescales and for different number of trades. The price return is defined as $\log m(t) - \log m(t - \tau)$, where τ is from left to right 1, 2, 4, 8, 16 and 32 market orders (continuous lines) and 5, 10, and 20 min (dashed lines). The data are for AZN, based on absolute returns (both buy and sell).
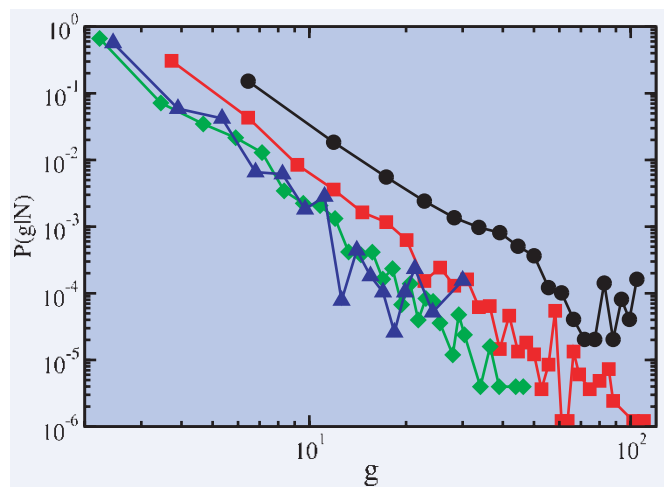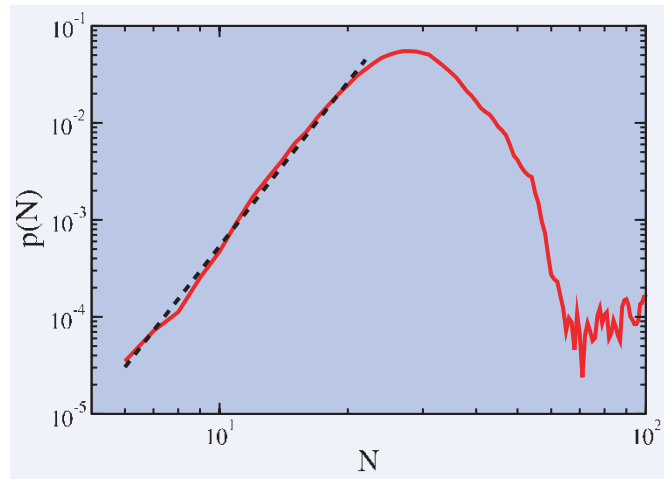
event scale are reflected on larger scales (at least as large as 20 min). Also, this figure makes it clear that the behaviour in real time is essentially the same as the behaviour in event time. We see similar behaviour for all the stocks in our sample. Since other studies have shown the continuity between timescales from 10 min to as long as a month [39, 47, 55], it seems clear that an explanation at the level of single events is sufficient.[18] This is particularly striking, given that, on a longer timescale, other processes may become important, such as order splitting. Figure 11 suggests that for understanding the statistical properties of prices these can be neglected.

The fact that we observe essentially identical distributions when we aggregate over individual transactions or over fixed time rules out any explanation of fat tails for price returns based on a subordinated random process, as was suggested by Clark [8]. While fluctuations in the number of events in a given length of time might be quite important for other phenomena, such as clustered volatility, they are clearly not important in determining the price return distribution.

7. Correlations in occupied sites

A natural hypothesis about the origin of the large gap distributions involves fluctuations in the number of occupied sites in the order book. If there are only a few orders in the book, so that it is mostly empty, we

Figure 12. The effect of a finite number of orders for the stock AZN. Panel (a) shows the unconditional distribution of the number of occupied price levels for buy orders, plotted on a double logarithmic scale. The dashed line is the best fit of the low N region with a functional form $p(N) \propto N^\beta$. Panel (b) shows the probability density of first gap sizes conditional on the number of occupied price levels, $p(g|N)$. We divide the sample into four subsamples according to the value of N. Specifically, we have $0 < N \leq 15$ (black circles), $15 < N \leq 30$ (red squares), $30 < N \leq 45$ (green diamonds), and $45 < N \leq 60$ (blue triangles). Surprisingly, the first gap size shows approximate power law behaviour even when the number of occupied sites is very large.

naturally expect large gaps to exist. Perhaps the approximate power for the gap distribution can be explained by a similar power law in the number of occupied sites? In this section we show that this is not the case, but rather that the power law behaviour depends on correlations in occupied price levels.

Interestingly, we do find evidence of power law scaling in the number of occupied price levels. Figure 12(a) shows the probability density function

---

[18] We have studied the response to individual events, and see some evidence for mean reversion of large price changes. However, the results of figure 11 make it clear that these are not sufficient to undo the initial change.

$p(N)$, where N is the number of occupied price levels in the order book at any given time. The low N region is well fit by a power law, $p(N) \sim N^{\beta}$, over about two decades.

However, the fluctuations in the number of orders stored in the book cannot explain the power law behaviour of prices, because the tail exponent of the power law behaviour of the gap distribution near zero is much too large. If we assume that the location of order deposition is uncorrelated, then an order arrival on a bounded domain gap size should be proportional to $1/N$. But with $p(N) \sim N^{\beta}$ and $g \sim 1/N$, $p(g) \sim g^{-(\beta+2)}$. For $\beta \approx 5.5$ this results in $p(g) \sim g^{-7.5}$. In contrast, the empirically observed scaling exponent for the density for AZN is about 3.5.

Surprisingly, the distribution of gap size shows approximate power law scaling, independent of N. This is illustrated for AZN in figure 12(b), where we plot the probability of gap size conditioned on N, $p(g|N)$, for several different values of N. The same approximate power law scaling behaviour is seen independent of N. The approximate power law behaviour is evident even with as many as 60 occupied sites on one side of the book. Considering that it is rare for orders to be placed more than 100 price ticks away from the best price, this illustrates that occupied sites display non-trivial correlations, which are essential for explaining large price fluctuations. This conclusion is reinforced by studies we have done comparing the real distribution of occupied levels with models based on IID order placement, which do not reproduce the power law behaviour.

8. Conclusions

For the London Stock Exchange, we have shown that large fluctuations in prices are unrelated to large transactions, or to the placement of large orders. Instead, large price fluctuations occur when there are gaps in the occupied price levels in the limit order book. Large changes occur when a market order removes all the volume at the best price, creating a change in the best price equal to the size of the gap.

At a higher level, these results demonstrate that large price changes are driven by fluctuations in liquidity. There are times when the market absorbs changes in supply and demand smoothly, and other times when a small change in supply or demand can result in a very large change in the price. This is due to the fact that supply and demand functions are not smooth, but rather have large, irregular steps and jumps. The market is granular, due to the presence of only a finite number of occupied price levels in the book. This is what in physics is called a finite size effect. Even for an active stock such as AZN, the number of occupied price levels on one side of the book at any given time is typically about 30, and so the system is far from the continuum limit.

However, we have shown that this is not a simple matter of fluctuations in the number of occupied price levels; while there is an approximate power law in the limit $N \to 0$ in the frequency for occupying N price levels, this is not sufficient to explain fluctuations in prices. Instead, the power law in gaps persists even when the number of occupied levels is quite high, reflecting non-trivial correlations in the positions of orders sitting in the book.

The empirical results that we have presented here raise as many questions as they answer. In particular, what is responsible for the approximate power law distribution of gap sizes, and why do limit orders have the same tail exponents as market orders and cancellations? We have done some modeling to address this question. In particular, we have modified the statistical model of order flow introduced by Daniels et al [13, 59] to include non-uniform order placement, more closely reflecting the empirical order placement distribution [6, 65], which has an approximate power law tail. For some parameter values our simulations reproduce the power law tails of the gap distribution (and hence of returns), but since we do not yet understand the necessary and sufficient conditions for this, we do not include these results here.

The work presented here suggests that it is important to properly model market institutions. However, at this point it is not clear how these results would change for a different market institution. Fat tails are also observed in markets, such as the London Metals Exchange, that follow very different exchange mechanisms.[19] While the LSE follows a continuous double auction, it is our belief that key elements are likely to persist with other market mechanisms. In particular, we hypothesize that large price fluctuations in any market are driven by liquidity fluctuations, and that the granularity of fluctuations in supply and demand remains the key factor underlying extreme price fluctuations.

These results are important because they reveal the detailed mechanism through which prices display (at least approximate) power law fluctuations. They suggest that many previous models that claimed to explain this phenomenon were misdirected, and provide strong constraints on future models. The tail exponent of large price changes appears to depend on parameters of the market: more lightly traded markets (with lower event rates) tend to display fatter tails, with more extreme risks. This has important practical importance because it gives some understanding of what determines financial risks, and gives some clues about how to reduce them.

---

[19] One of us (JDF) did an unpublished study in 1995 of daily data for industrial metals such as copper, zinc, iron, nickel, etc. traded on the London Metals Exchange, whose exchange mechanism resembles a Walrasian market. All of these have large kurtosis, indicating fat tails.

Acknowledgements

References

[1] Akgiray V, Booth G G and Loistl O 1989 Stable laws are inappropriate for describing German stock returns *Allegemeines Statistisches* **73** 115–21

[2] Anteneodo C and Tsallis C 2004 Multiplicative noise: a mechanism leading to nonextensive statistical mechanics *J. Math. Phys.* forthcoming

[3] Arthur W B, Holland J H, LeBaron B, Palmer R and Tayler P 1997 Asset pricing under endogenous expectations in an artificial stock market *The Economy as an Evolving Complex System II* eds W B Arthur, S N Durlauf and D H Lane (Redwood City: Addison-Wesley) pp 15–44

[4] Biais B, Hillion P and Spatt C 1995 An empirical analysis of the limit order book and order flow in the Paris bourse *J. Finance* **50** 1655–89

[5] Bouchaud J-P, Gefen Y, Potters M and Wyart M 2004 Fluctuations and response in financial markets: the subtle nature of 'random' price changes *Quant. Finance* **4** 176–90

[6] Bouchaud J-P, Mezard M and Potters M 2002 Statistical properties of the stock order books: empirical results and models *Quant. Finance* **2** 251–6

[7] Challet D, Chessa A, Marsili M and Zhang Y C 2001 From minority games to real markets *Quant. Finance* **1** 168–76

[8] Clark P K 1973 Subordinated stochastic process model with finite variance for speculative prices *Econometrica* **41** 135–55

[9] Cohen K J, Maier S F, Schwartz R A and Whitcomb D K 1981 Transaction costs, order placement strategy, and existence of the bid–ask spread *J. Political Economy* **89** 287–305

[10] Cont R and Bouchaud J-P 2000 Herd behavior and aggregate fluctuations in financial markets *Macroeconomic Dynamics* **4** 170–96

[11] Coppejans M and Domowitz I 2002 An empirical analysis of trades, orders, and cancellations in a limit order market *Economics Department Working Paper* Duke University

[12] Coppejans M, Domowitz I and Madhavan A 2003 Dynamics of liquidity in an electronic limit order book market *Economics Department Working Paper* Duke University

[13] Daniels M G, Farmer J D, Gillemot L, Iori G and Smith E 2003 Quantitative model of price diffusion and market friction based on trading as a mechanistic random process *Phys. Rev. Lett.* **90**

[14] Evans M D D and Lyons R K 2002 Order flow and exchange rate dynamics *J. Political Economy* **110** 170–80

[15] Fama E F 1965 The behavior of stock-market prices *J. Business* **38** 34–105

[16] Farmer J D 1996 Slippage 1996 *Prediction Company Technical Report*, Santa Fe (http://predict.com/jdf/slippage.pdf)

[17] Farmer J D and Lillo F 2004 On the origin of power laws in financial markets *Quant. Finance* **4** 7–10

[18] Farmer J D, Patelli P and Zovko I 2003 The predictive power of zero intelligence (http:xxx.lanl.gov/cond-mat/0309233)

[19] Foucault T 1999 Order flow composition and trading costs in a dynamic limit order market *J. Financial Markets* **2** 99–134

[20] Gabaix X, Gopikrishnan P, Plerou V and Stanley H E 2003 A theory of power-law distributions in financial market fluctuations *Nature* **423** 267–70

[21] Ghashghaie S, Breymann W, Peinke J, Talkner P and Dodge Y 1996 Turbulent cascades in foreign exchange markets *Nature* **381** 767–70

[22] Giardina I and Bouchaud J-P 2003 Bubbles, crashes and intermittency in agent based market models *Eur. Phys. J.* B **31** 421–37

[23] Goldstein M A and Kavajecz K A 2000 Eights, sixteenths, and market depth: changes in tick size and liquidity provision on the NYSE *J. Financial Economics* **56** 125–49

[24] Gopikrishnan P, Plerou V, Gabaix X and Stanley H E 2000 Statistical properties of share volume traded in financial markets *Phys. Rev.* E **62** R4493–96

[25] Griffiths M D, Smith B F, Turnbill D A S and White R W 2000 The costs and determinants of order aggressiveness *J. Financial Economics* **56** 65–88

[26] Hasbrouck J 1991 Measuring the information content of stock trades *J. Finance* **46** 179–207

[27] Hausman J A, Lo A W and Mackinlay A C 1992 An ordered probit analysis of transaction stock prices *J. Financial Economics* **31** 319–79

[28] Hill B M 1975 A simple general approach to inference about the tail of distribution *Ann. Statistics* **3** 1163–74

[29] Hollifield B, Miller R A and Sandas P 2002 Empirical analysis of limit order markets *Working Paper* Carnegie Mellon University *Rev. Economic Studies* to appear

[30] Johnson N F, Jeffries P and Hui P M 2003 *Financial Market Complexity* (Oxford: Oxford University Press)

[31] Kaniel R and Liu H 2003 So what do informed traders use? *Working Paper* Duke University

[32] Kavajecz K A 1999 A specialist's quoted depth and the limit order book *J. Finance* **54** 747–71

[33] Kavajecz K A and Odders-White E R 2003 Technical analysis and liquidity provision *Working Paper* University of Wisconsin

[34] Kempf A and Korn O 1999 Market depth and order size *J. Financial Markets* **2** 29–48

[35] Knez P J and Ready M J 1996 Estimating the profits from trading strategies *Rev. Financial Studies* **9** 1121–63

[36] Koedijk K G, Schafgans M M A and de Vries C G 1990 The tail index of exchange rates *J. Int. Economics* **29** 1–197

[37] Lillo F, Farmer J D and Mantegna R N 2003 Econophysics—master curve for price–impact function *Nature* **421** 129–30

[38] Lillo F and Mantegna R N 2003 Power-law relaxation in a complex system: Omori law after a financial market crash *Phys. Rev.* E **68**

[39] Longin F M 1996 The asymptotic distribution of extreme stock market returns *J. Business* **69** 383–408

[40] Loretan M and Phillips P C B 1994 Testing the covariance stationarity of heavy-tailed time series: an overview of the theory with applications to several financial datasets *J. Empirical Finance* **1** 211–48

[41] Lux T 1996 The stable Paretian hypothesis and the frequency of large returns: an examination of major German stocks *Appl. Financial Economics* **6** 463–75

[42] Lux T and Marchesi M 1999 Scaling and criticality in a stochastic multi-agent model of a financial market *Nature* **397** 498–500

[43] Mandelbrot B 1963 The variation of certain speculative prices *J. Business* **36** 394–419

[44] Mandelbrot B 1997 *Fractals and Scaling in Finance* (New York: Springer)

[45] Mandelbrot B and Taylor H 1967 On the distribution of stock prices differences *Operations Research* **15** 1057–62

[46] Mantegna R N and Stanley H E 1995 Scaling behavior in the dynamics of an economic index *Nature* **376** 46–9

[47] Mantegna R N and Stanley H E 1999 *Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge: Cambridge University Press)

[48] Muller U A, Dacorogna M M and Pictet O V 1998 Heavy tails in high-frequency financial data *A Practical Guide to Heavy Tails: Statistical Techniques and Applications* eds R J Adler, R E Feldman and M S Taqqu (Berlin: Springer) pp 55–78

[49] Niederhoffer V and Osborne M F M 1966 Market making and reversal on the stock exchange *J. Am. Statistical Assoc.* **61** 897–916

[50] Officer R R 1972 Distribution of stock returns *J. Am. Statistical Assoc.* **67** 807–12

[51] Osler C L 2003 Currency orders and exchange rate dynamics: an explanation for the predictive success of technical analysis *J. Finance* **58** 1791–1819

[52] Parlour C A 1998 Price dynamics in limit order markets *Rev. Financial Studies* **11** 789–816

[53] Petersen M A and Umlauf S 1994 Quote revisions and the quoted size: an empirical analysis *Kellogg School of Management Working Paper Series*, Northwestern University

[54] Plerou V, Gopikrishnan P, Amaral L A N, Gabaix X and Stanley H E 2000 Economic fluctuations and anomalous diffusion *Phys. Rev.* E **62** R3023–26

[55] Plerou V, Gopikrishnan P, Amaral L A N, Meyer M and Stanley H E 1999 Scaling of the distribution of price fluctuations of individual companies *Phys. Rev.* E **60** 6519–29

[56] Plerou V, Gopikrishnan P, Gabaix X and Stanley H E 2002 Quantifying stock price response to demand fluctuations *Phys. Rev.* E **66**

[57] Potters M and Bouchaud J-P 2003 More statistical properties of order books and price impact *Physica* A **324** 133–40

[58] Rosenow B 2002 Fluctuations and market friction in financial trading *Int. J. Mod. Phys.* C **13** 419–25

[59] Smith E, Farmer J D, Gillemot L and Krishnamurthy S 2003 Statistical theory of the continuous double auction *Quant. Finance* **3** 481–514

[60] Sornette D 1998 Multiplicative processes and power laws *Phys. Rev.* E **57** 4811–3

[61] Sornette D 2000 *Critical Phenomena in Natural Sciences* (Berlin: Springer)

[62] Takayasu H 2002 *Empirical Science of Financial Fluctuations* (Berlin: Springer)

[63] Torre N 1997 *BARRA Market Impact Model Handbook* (Berkeley: BARRA Inc)

[64] Weber P and Rosenow B 2004 Large stock price changes: volume or liquidity? (http://xxx.lanl.gov/cond-mat/041132)

[65] Zovko I and Farmer J D 2002 The power of patience; a behavioral regularity in limit order placement *Quant. Finance* **2** 387–92