# On the origin of power-law fluctuations in stock prices

*Vasiliki Plerou*, *Parameswaran Gopikrishnan*, *Xavier Gabaix* and *H Eugene Stanley* respond to the comments on their recent article by Farmer and Lillo.

## Introduction

We recently proposed a testable theory for the origin of the empirically-observed power-law distributions of financial market variables such as stock returns, volumes and frequency of trades [1]. Our theory explains the power-law exponent of the distribution of returns by deriving a square-root functional form for market impact ('square-root law') that relates price impact and order size. Our previous empirical analysis gave results that support the square-root law of market impact.

Farmer and Lillo (FL) [2] raise some issues related to the empirical validity of the theory proposed in [1]. Their discussion is based on the following arguments:

1. FL claim that the price impact function grows slower than a square-root law. Interestingly, FL's empirical analysis does find a power-law relationship for market impact with the exponent $\beta \approx 0.5$ for volumes smaller than a threshold, which is consistent with the square-root form of market impact given by our theory. However, for large volumes, FL claim that $\beta < 0.2$ (for the New York Stock Exchange—NYSE) and $\beta \approx 0.26$ (based on analysing three stocks in the London Stock Exchange—LSE). FL do not compute error bars for $\beta$ for large volumes, but claim that, from a visual comparison, $\beta = 0.5$ (square-root law) is inconsistent with the data.
2. FL argue that the empirical analysis that we presented in support of the square-root functional form of market impact [2] is 'invalidated' by the 'long-memory nature of order flow'.
3. FL analyse the volume distribution of three stocks from the LSE and claim that the volume distribution does not follow a power law. Consequently, FL conclude that volume fluctuations do not determine the power-law tail of returns.

We reply to these issues by first addressing these criticisms and then presenting our response with results of our new analysis.

1. FL find $\beta < 0.5$ for large volumes from analysing the average value of return for a trade for a given trade size. FL's procedure for estimating price impact is flawed since large orders are usually executed by splitting into smaller-sized trades, so the procedure used by FL gives a downward bias for the power-law exponent $\beta$ defined in our theory [1, 5], giving rise to an apparent exponent value $\beta'$ smaller than the correct value $\beta$. In fact FL's procedure gives $\beta < 0.5$ for large volumes—precisely the domain in which we expect the order-splitting effect to be dominant—and therefore a downward bias for $\beta$.
2. Although we present new estimators to address this point, we believe FL's argument to be incorrect since long-memory in order flow clearly does not imply the same for

returns, so FL's criticisms about our estimation procedure do not seem relevent. To address a potential problem of long memory in order flow, we draw from a new estimator for measuring market impact [3] and extend our previous analysis to the 1000 largest NYSE stocks. Our new estimation confirms that the market impact function does behave as a square-root function of the volume.

3. We analyse 262 largest stocks listed in the LSE. Our analysis of volume distribution for these 262 stocks shows that the distribution of volume decays as a power law with an exponent $\approx 1.5$ in agreement with our previous results for the NYSE and the Paris Bourse. In fact our analysis of the volume distribution for the same stocks analysed by FL shows a clear power-law behaviour with exponent $\approx 1.5$—in contrast to FL's claim.

Define $S_i$ as the price of the stock after trade $i$,

$$\delta p_i \equiv \log S_i - \log S_{i-1} \tag{1}$$

as the return concomitant to trade $i$, so the return over a fixed time interval $\Delta t$ is

$$r \equiv \sum_{i=1}^{N} \delta p_i, \tag{2}$$

where $N$ is the number of trades in $\Delta t$. Let $q_i$ be the number of shares traded in trade $i$, so

$$Q \equiv \sum_{i=1}^{N} q_i \tag{3}$$

is the total volume in interval $\Delta t$. We define the trade imbalance

$$\Omega \equiv \sum_{i=1}^{N} \epsilon_i q_i \tag{4}$$

where $\epsilon_i = 1$ indicates a buyer-initiated trade and $\epsilon = -1$ denotes a seller-initiated trade. We denote by $V$ the size of a large order, which can be executed in several trades.

## 1. Measuring market impact

Let $\Delta p$ be the change in price caused by a large order of size $V$; all else remains the same. Our theoretical approach [1] derives a power-law functional form for the market impact function[1],

$$\Delta p \sim V^{\beta}. \tag{5}$$

We hypothesized $\beta = 0.5$ and supported this using empirical analysis.

Our hypothesis equation (5) pertains to $\Delta p$, the total impact in price of a large order of size $V$. In practice, as in [1], large orders are executed by splitting into orders of smaller size which are observed in the trade time series as the trade size $q_i$. The empirical analysis of [2] and [4] refer to the relationship of local price change $\mathbf{E}(\delta p|q)$ and not the price impact $\mathbf{E}(\Delta p|V)$ that we are interested in. The true market impact function $\mathbf{E}(\Delta p|V)$ is indeed notoriously difficult to measure since the information about the unsplit order size is usually proprietary and not available, either in our data or in the data analysed by FL.

FL claim that the price impact function grows more slowly than a square-root function for large volumes. The basis of their claim is the analysis in [4] that $\mathbf{E}(\delta p|q) \sim q^{\beta}$ with $\beta = 0.5$ for small $q$ and $\beta = 0.2$ for larger $q$. While $\mathbf{E}(\delta p|q)$ indeed grows less rapidly than a square root, as reported in [10], $\mathbf{E}(\delta p|q)$ neither quantifies price impact of large trades, nor does it contradict our theory and empirical results [1]. This is because a trade-by-trade analysis of $\mathbf{E}(\delta p|q)$ leads to a *biased* measurement of full price impact and the exponent $\beta$, since it does not take into account the splitting of trades [1, 5].

Consider an example. Suppose that a large fund wants to buy a large number $V$ of shares of a stock whose price is \$100. The fund's dealer may offer this large volume for a price of \$101. Before this transaction, however, the dealer must buy the shares. The dealer will often do that progressively in many steps, say 10 in this example. In the first step, the dealer will buy $V/10$ shares, and the price will go say, from \$100 to \$100.1, and in the second the price will go from \$100.1 to \$100.2. After some time elapses, the price will have gone to \$101 in increments of \$0.1. At this stage, the dealer has his required number of shares, and hands them over to the fund manager at a price of \$101. The true price impact here is 1%, since the price has gone from \$100 to \$101. But in any given transaction, the price has moved by no more than \$0.1. So [2] would find an 'apparent' price impact of no more than \$0.1, i.e. 0.1% of the price. Since as the transaction is executed the price of the stocks goes from \$100 to \$101, the true price impact is 1%. As a result the procedure of FL will measure a value 10 times smaller than the true value. This downward bias explains why FL find in figure 2 a maximum impact of 0.1%—a very small price impact. Other evidence in economics [7, 8] finds impacts that are up to 40 times larger than that of FL's analysis. Likewise our evidence pertains to large impacts, captured by figure 2 of [1] which shows on the vertical axis values of $r^2$ equal up to 200 times the variance $\sigma^2$, i.e. values of return $r$ up to $\sqrt{200} \approx 14$ standard deviations.

We can quantify the bias in the above example. Suppose that a trade of size $V$ is split into $K = V^{\alpha}$ (10 in our example) trades of equal size $q = V/K = V^{1-\alpha}$, with $0 < \alpha < 1$. Then the apparent impact $\delta p$ incurred by each trade (0.1 % in our example) will be $1/K$ (1/10 in our example) of the total price impact $V^{\beta}$ (1% in our example), i.e. $\delta p = V^{\beta}/K = V^{\beta-\alpha}$. So a power-law fit of $\delta p$ versus $q$, such as the one presented in figure 2 of FL, will give $\delta p \sim q^{\beta'}$ with[2]

$$\beta' = (\beta - \alpha)/(1 - \alpha) < \beta.$$

The 'trade-by-trade' measurement of the price impact, as performed by [2, 4], leads to a biased measurement $\beta'$ of the exponent $\beta$ of the true price impact.

[1] As in [1], we interpret $\Delta p \sim V^{\beta}$ to mean that there exist a slowly varying function $L(V)$ such that for large $V$, $\Delta p / V^{\beta} L(V) \to 1$. A function is called slowly varying when for all $t > 0$, $\lim_{V\to\infty} L(tV)/L(V) = 1$. Typical slowly varying functions are $L(V) = a$, or logarithmic corrections: $L(V) = a \ln(V)^{\alpha}$, where $a$ and $\alpha$ are constants.

[2] The inequality below holds if $\beta < 1$. There is wide agreement that $\beta$ should be no greater than 0.5. This is because the power-law exponent of returns $\zeta_r$ and the power-law exponent of volumes satisfy $\zeta_r \leqslant \zeta_V/\beta$ [1], so that the empirical values $\zeta_V \approx \zeta_r/2 \approx 1.5$ imply $\beta \precsim 0.5$.

It is to address this bias that we examine $\mathbf{E}(r^2|Q)$ in [1]. As is well established empirically, the sign of returns is unpredictable in the short term, so the reasoning in [5] shows that $\mathbf{E}(r^2|Q)$ will not be biased[3].

Our analysis [1] was presented with data for the 116 most actively traded stocks. To check if the result of $\beta = 0.2$ for large volumes presented in [2] and [4] could arise from increasing the size of the database, we now extend our analysis to the 1000 largest stocks in our database for the 2-year period 1994–1995. Figure 1(*a*) confirms that $\mathbf{E}(r^2|Q) \sim Q$ as predicted by our theory.

## 2. Robustness of estimation against the long memory of order flow

FL argue that the empirical analysis that we presented in support of the square-root functional form of market impact is 'invalidated' by the 'long-memory nature of order flow'. FL's argument is based entirely on the *assumption* that returns due to each transaction $i$ can be written as $r_i = \epsilon_i q_i^\beta$ where $\epsilon_i = 1$ for a buy trade and $\epsilon_i = -1$ for a sell trade. Under this assumption, FL then argue that our estimator $\mathbf{E}(r^2|Q)$ is affected by the long-range correlations in the trade signs $\epsilon_i$ [2, 11]. FL give some numerical evidence for this potential effect for small to moderate volumes.

All of the tests shown by FL are for a *fictitious* return $f_i$ constructed on a trade-by-trade basis as $f_i = \epsilon_i q_i^\beta$. FL's argument and the tests shown in figure 1 of FL are for the fictitious return. In reality, returns certainly cannot be expressed as $r_i = \epsilon_i q_i^\beta$, with $\epsilon_i$ being the trade indicator. Indeed, if this were true, returns themselves would be long-range correlated—a possibility long known to be at odds with empirical data. Since the sign of the return $r_i$ and that of the trade sign $\epsilon_i$ are clearly not equal, FL's argument about our estimation procedure being affected by the long-memory nature of the trade sign ($\epsilon_i$) is incorrect. (See [11] on a related point.)

Although FL's argument is incorrect, to address the general concern that the autocorrelations of the trade signs $\epsilon_i$ might bias our analysis, we draw from a forthcoming paper [3], which performs the following analysis[4]. For each interval $\Delta t$ define $q_{\max}$ as the size of the largest trade. In our theory, if the largest trade $V_{\max}$ is large, it will have a the major influence on the value of return, so that one will have $r^2 \sim V_{\max}^{2\beta} \sim q_{\max}^{2\beta}$. Hence $q_{\max}$ gives us a diagnostic value of the behaviour of the largest trade, independently of a potential collective behaviour. We detail this in [3]. We compute $\mathbf{E}(r^2|q_{\max})$ and find (figure 1(*b*))

$$\mathbf{E}(r^2|q_{\max}) \sim q_{\max}. \qquad (6)$$

[3] Our theory is one of *large* trades. The evidence presented in [2] concerns small to moderate size trades. It follows a long tradition pioneered by [6] (see also [10, 11]), and may be fine for small to moderate trades. The issue of splitting may be less important for those trades, but it is almost certainly crucial for large ones (e.g. trades bigger than 5 to 10 times the average trade size). This is why the trade-by-trade $\mathbf{E}(\delta p|q)$ measurement has a downward bias for large trade, which is avoided when we take $\mathbf{E}(r^2|Q)$.

[4] We give the basis of this analysis in [3]. The specifics of the split matter in principle. In [1] we present a theory of power-law splitting in which the size of the largest chunk, $q_{\max}$, is proportional to that of the entire order, $V$.

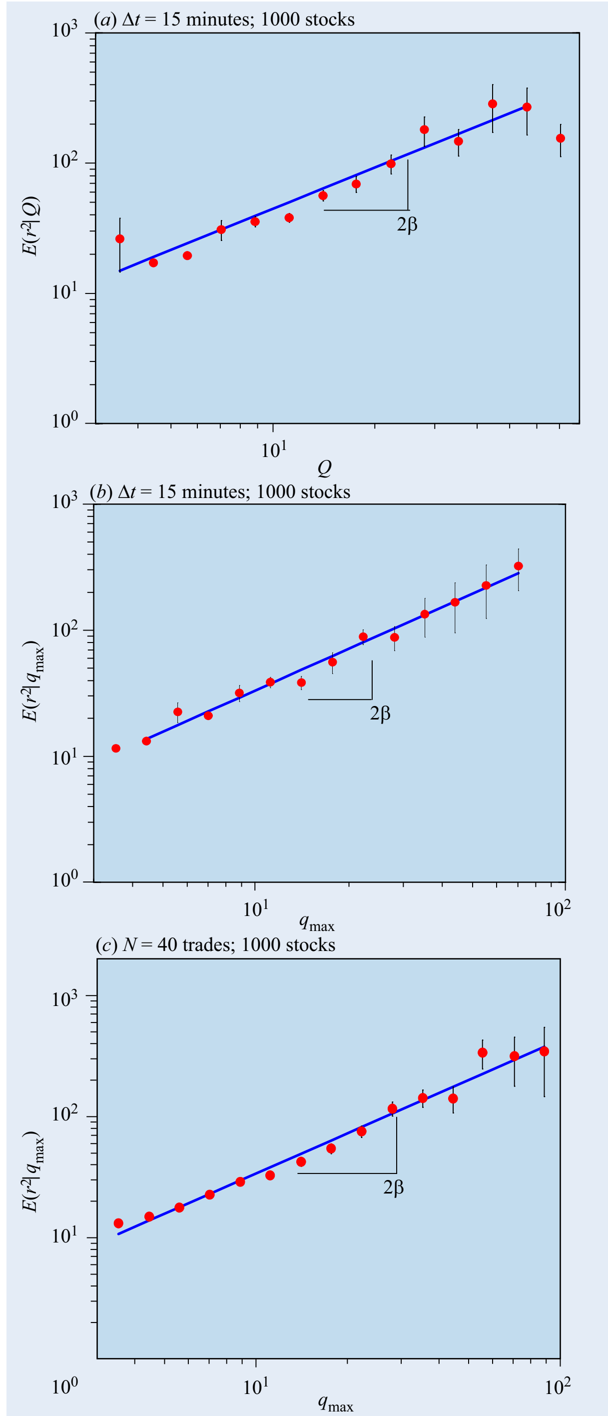


**Figure 1.** (*a*) Conditional expectation function $\mathbf{E}(r^2|Q)$ of the squared return for a given volume for $\Delta t = 15$ minutes for 1000 largest stocks in the NYSE, Amex and NASDAQ (1994–1995). We normalized $r$ for each stock to zero mean and unit variance, and normalized $Q$ by its first centred moment. This normalization procedure allows for a data collapse for different stocks and the plot represents the average for 1000 stocks. Regressions in the range $3 < Q < 70$ give values of $\beta = 1.05 \pm 0.08$. (*b*) $\mathbf{E}\left(r^2|q_{\max}\right)$ of the squared return for a given $q_{\max}$ for $\Delta t = 15$ minutes. Here $q_{\max}$ is the largest trade size in the 15 minute interval. Power-law regression gives the value of $2\beta = 1.09 \pm 0.06$ consistent with $\beta = 0.5$. (*c*) $\mathbf{E}\left(r^2|q_{\max}\right)$ of the squared return for a given $q_{\max}$ for fixed number of trades $N = 40$. Source: [3].

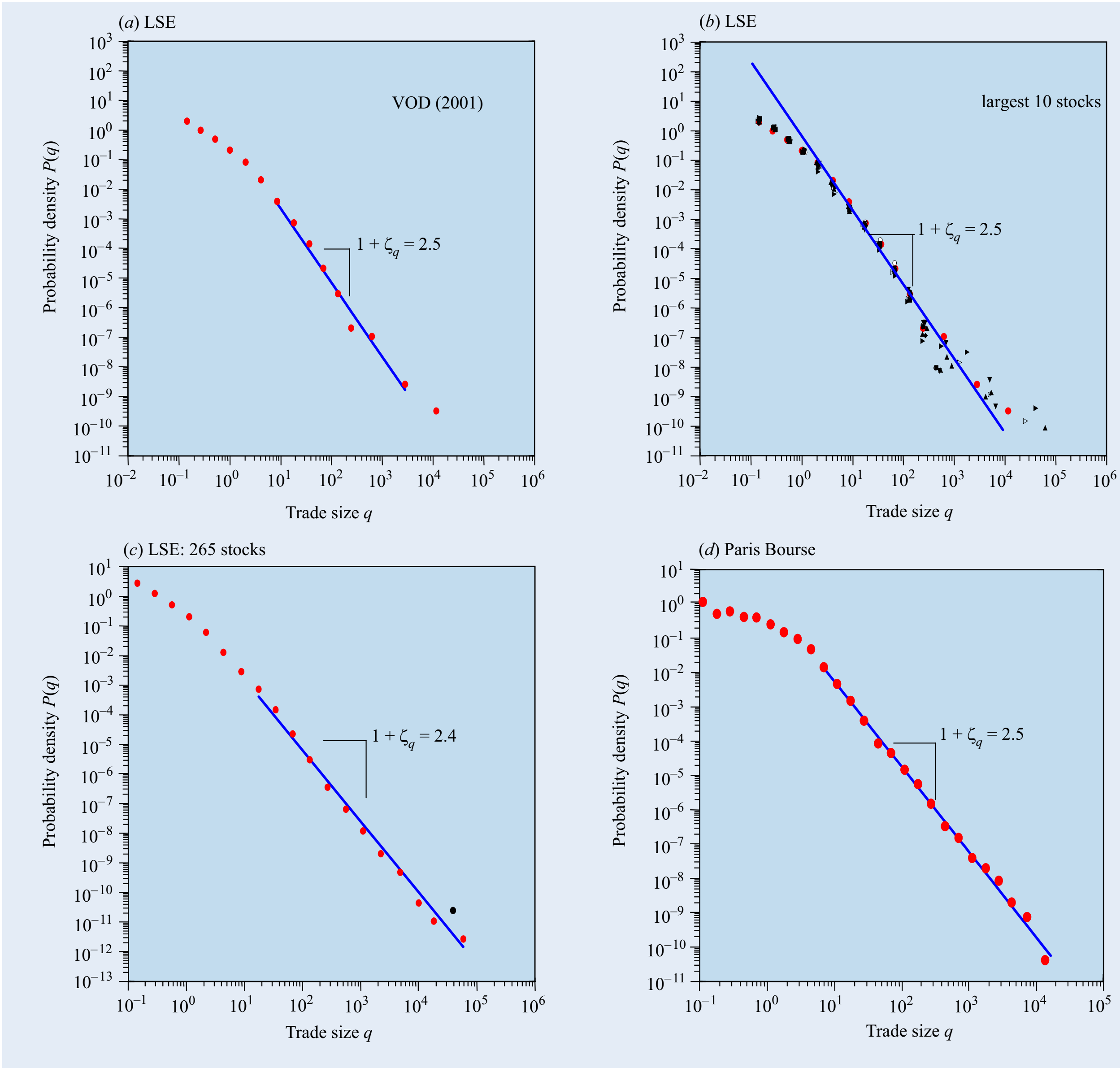


**Figure 2.** (*a*) Probability density function of trade volumes for Vodafone (VOD) for 2001. A power-law fit in the region $10 < q < 1000$ gives a value of the exponent $\zeta_q = 1.5 \pm 0.1$. In contrast FL finds much more rapid decay. (*b*) Probability density function of trade volumes for 10 largest stocks listed in the LSE, showing clear evidence of power-law decay with exponent $\zeta_q = 1.5 \pm 0.1$, consistent with our previous results [9] for the NYSE and for the Paris Bourse. Here $q$ are normalized by its first centred moment, so all 10 distributions collapse on one curve. We find an average exponent $\zeta_q$ =1.59 ±0.09. (*c*) Same as (*b*) but for all 265 stocks in our sample where $q$ is normalized for each stock by its first centred moment. We find $\zeta_q = 1.4 \pm 0.09$. (*d*) Probability density function of trade volumes for 30 largest stocks listed in the Paris Bourse obtained by the same procedure. We find $\zeta_q = 1.49 \pm 0.03$ [5].

In addition to the above, to ensure that our estimation is robust to varying number of trades in a fixed $\Delta t$, we have computed $\mathbf{E}(r^2|q_{\max})$ for fixed number of trades instead. Figure 1(*c*) shows that $\mathbf{E}(r^2|q_{\max}) \sim q_{\max}$ for $r$ over $N = 40$ trades. It can be seen that a power-law regression gives the value of $2\beta = 1.10 \pm 0.06$.

We would like to emphasize that in figure 1 we consider very large trades that are up to 70 times the first moment of volume. They correspond to returns of up to 14 standard deviations of returns. This confirms that we study very large trades and returns—the ones that are relevant for the study of power-law fluctuations, while in contrast FL's analysis does not systematically treat large trades.

We conclude that the procedure used in [2] has a downward bias of the price impact $\beta$ of large trades. When we perform more appropriate analysis, we confirm that the $\beta \approx 0.5$. This corroborates our hypothesis [1, 5] that large fluctuations in volume cause large fluctuations of prices.

### 3. Half-cubic power-law distribution of volumes

The last claim of FL pertains to the very nature of the volume distribution. They present the results of their analysis of three stocks in the LSE and claim their analysis shows no evidence for a power-law distribution.

We analyse the same database which records *all* trades for *all* stocks listed in the LSE. From this database, we first examine one stock—Vodafone, VOD—which is analysed by FL. For this stock, we compute the volume distribution and find clear evidence for a power-law decay (figure 2(*a*))

$$P(q) \sim q^{-\zeta_q - 1} \quad (7)$$

with $\zeta_q = 1.5 \pm 0.1$, in agreement with our results for the NYSE and the Paris Bourse [1,9], but in sharp contrast to the FL results who claim a thin-tailed distribution for the same data.

For the 10 largest stocks in our sample, figure 2(*b*) shows that $P(q)$ is consistent with the same power law of $\zeta_q = 1.5$ which is consistent with our earlier finding for the NYSE [9]. We extend our analysis to the 250 largest stocks and find similar results (see figure 2(*c*)).

To test the universal nature of this distribution, we analyse data for 30 largest stocks listed in the Paris Bourse and find that $P(q)$ is consistent with the a power law with almost identical exponents $\zeta_q = 1.5$ (figure 2(*d*)).

In summary, the analysis of [2] pertains to small to moderate trades. FL's estimation is biased for large trades, so FL can detect only very small price impacts, less than 0.1%. When we use our more general procedure and study significantly larger data, we confirm our initial finding of a square root price impact function. We conclude that the available evidence is consistent with our hypothesis [1, 5] that large fluctuations the volume traded by large market participants may contribute significantly to the large fluctuations in stock prices [3].

### *Note added in proof*

After [2], it has become clear that FL's claim of a non-power-law distribution of trade sizes for the LSE stocks is based on incomplete data. FL's analysis excludes the upstairs market[5] trades which contain the largest trades in the LSE. In contrast, our result of a 1.5 power-law exponent for the volume distribution is based on data containing all trades (both the upstairs and the downstairs market trades) in the LSE. By excluding the large trades in the upstairs market, FL set an artificial truncation at large volume, so FL's finding of a non-power-law distribution of volume is a trivial artifact of incomplete data. Although FL claim in their note added in proof that 'it has been shown that large price fluctuations in the NYSE (including the upstairs market) and the electronic portion of the LSE are driven by fluctuations in liquidity' their new analysis and findings are affected by the same problems as in their present comment: (i) incompleteness (absence of the upstairs market trades) of the data analysed and (ii) they do not take into account the splitting of large orders.

Gabaix *et al* [1, 5] and FL [2] discuss two distinct possibilities respectively: (i) large price changes arise from large trades and (ii) large price changes arise from fluctuations in liquidity [10, 12]. While we believe that both mechanisms play a role in determining the statistics of price changes, our empirical findings support the possibility that the specific power-law form of the return distribution arises from large trades.

[5] Large trades tend to be executed in the 'upstairs' market by bilateral arrangements than through the order book.

### Acknowledgments

We thank the NSF's economics programme for financial support, and J-P Bouchaud, Joel Hasbrouck, Carl Hopman, Gideon Saar, Jeff Wurgler and especially Gilles Zumbach.

### References

[1] 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

[2] Farmer J D and Lillo F 2004 On the origin of power-law tails in price fluctuations *Quant. Finance* **4** C7–11

[3] Plerou V, Gopikrishnan P, Gabaix X and Stanley H E 2004 On the relation between very large trades and very large price movements, in preparation

[4] Lillo F, Farmer J D and Mantegna R N 2003 Master curve for price-impact function *Nature* **421** 129–30

[5] Gabaix X, Gopikrishnan P, Plerou V and Stanley H E 2004 A theory of large movements in stock market activity *Working paper* http://ssrn.com/abstract=442940

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

[7] Chan L and Lakonishok J 1995 The behavior of stock prices around institutional trades *J. Finance* **50** 1147–74

[8] Keim D and Madhavan A 1997 Transactions costs and investment style. An inter-exchange analysis of institutional equity trades. *J. Financial Economics* **46** 265–92

[9] Gopikrishnan P, Plerou V, Gabaix X and Stanley H E 2000 Statistical properties of share volume traded in financial markets *Phys. Rev.* E **62** R4493–6

[10] Plerou V, Gopikrishnan P, Gabaix X and Stanley H E 2002 Quantifying stock price response to demand fluctuations *Phys. Rev.* E **66** 027104

[11] 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

[12] Plerou V, Gopikrishnan P, Amaral L A N, Gabaix X and Stanley H E 2000 Economic fluctuations and anomalous diffusion *Phys. Rev.* E **62** R3023–6

*Vasiliki Plerou and H Eugene Stanley are at the Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215, USA*

*Xavier Gabaix is at the Economics Department, MIT and NBER, Cambridge, MA 02142, USA*

*Parameswaran Gopikrishnan is at Goldman Sachs & Co., New York, NY 10004, USA.*