On the origin of power-law tails in price fluctuations

J Doyne Farmer and *Fabrizio Lillo* discuss the theory presented by Gabaix *et al* in their recent article published in *Nature* and offer an alternative analysis.

Introduction

Gabaix *et al* [1] recently proposed a testable theory for the origin of power-law tails in price fluctuations. In essence, their proposal is that they are driven by fluctuations in the volume of transactions which have a power-law tail with an exponent of roughly -1.5. These are modulated by a deterministic market impact function which describes the response of prices to transactions. More specifically, they argue that the distribution of large trade sizes scales as $P(V > x) \sim x^{-\gamma}$, where V is the volume of the trade and $\gamma \approx 1.5$. Based on the assumption that agents are profit optimizers, they argue that the average market impact function[1] is a deterministic function of the form $r = kV^{\beta}$, where r is the the change in the logarithm of price resulting from a transaction of volume V, k is a constant and $\beta = 0.5$. This implies that large price returns r have a power-law distribution with exponent $\alpha = \gamma/\beta \approx 3$. They argue that their theory is consistent with the data, even though these results are inconsistent with several previous studies [2–4] in the same markets they study (the New York Stock Exchange (NYSE) and Paris Stock Exchange (PSE)).

1. Problems with the test of Gabaix *et al*

Gabaix *et al* [1] present statistical evidence that appears to show that the NYSE and PSE data are consistent with the hypothesis that the average market impact function follows a square-root law which gives power-law tails for prices with an exponent of roughly -3. We show that their test fails to take into account the long-memory properties of the data. This dramatically weakens their test, so that it lacks the power to reject reasonable alternative hypotheses and gives misleading results.

Here we present an analysis that properly takes long-term memory into account. This shows that the functional form of the average market impact function varies from market to market, and in some cases from stock to stock. In fact, for both the London Stock Exchange (LSE) and the NYSE the average market impact function grows much slower than a square-root law; this implies that the exponent for price fluctuations predicted by modulations of volume fluctuations is much too big. We find that for LSE stocks the distribution of transaction volumes does not even have a power-law tail, thus volume fluctuations do not determine the power-law tail of price returns.

[1] One should more properly think of the market impact as a response to the order initiating the trade. That is, in every transaction there is a 'just-arrived' order that causes the trade to happen, and this order tends to alter the best quoted price in the direction of the trade, e.g. a buy order tends to drive the price up, and a sell order tends to drive it down.

Gabaix *et al*'s method to test the hypothesis of square-root price impact is to investigate $E[r^2|V]$ over a given time interval, e.g. 15 minutes, where r is the price shift and $V = \sum_{i=1}^{M} V_i$ is the sum of the volumes of the M transactions occurring in that interval. They chose to analyse r^2 rather than r because of its properties under time aggregation. To see why this might be useful, assume the return due to each transaction i is of the form $r_i = k\epsilon_i V_i^{\beta} + u_i$, where u is an IID noise process that is uncorrelated with V_i and ϵ_i is the sign of the transaction. The squared return for the interval is then of the form

$$r^2 = \sum_{i=1,j=1}^{M} \left(k\epsilon_i V_i^{\beta} + u_i\right)\left(k\epsilon_j V_j^{\beta} + u_j\right). \quad (1)$$

Under the assumption that V_i, V_j, ϵ_i and ϵ_j are all uncorrelated, when $\beta = 0.5$ it is easy to show that $E[r^2|V] = a + b\ V$, where a and b are constants.

The problem is that for the real data V_i, V_j, ϵ_i and ϵ_j are strongly correlated and, indeed, the sequence of signs ϵ_i is a long-memory process [5, 6]. To demonstrate the gravity of this problem, we use real transactions V_i, but introduce an artificial and deterministic market impact function of the form $r_i = kV_i^{\beta}$ with $\beta \neq 0.5$. We first fix the number of transactions, and then repeat the same procedure using a fixed time period. We examine blocks of trades with M transactions, $\{\epsilon_i, V_i\}$, $i = 1, \ldots, M$, where $\epsilon_i = +1$ (-1) for buyer (seller) initiated trades and V_i is the volume of the trade in number of shares. For each trade we create an artificial price return $r_i = k\epsilon_i V_i^{\beta}$, where k is a constant. Then for each block of M trades we compute $r = \sum_{i=1}^{M} r_i = k \sum_{i=1}^{M} \epsilon_i V_i^{\beta}$ and $V = \sum_{i=1}^{M} V_i$. Since we are using the real order flow we are incorporating the correct autocorrelation of the signs ϵ_i and transaction sizes V_i. Figure 1(*a*) shows $E[r^2|V]$ for different values of M and $\beta = 0.3$ for the British stock Vodafone in the period from May 2000 to December 2002, a series which contains approximately 10^6 trades.

For small values of M the quantity $E[r^2|V]$ follows the artificial market impact functional form $E[r^2|V] \sim V^{2\beta} = V^{0.6}$, but when M is large the relation between $E[r^2|V]$ and V becomes linear. The value $M = 40$ is roughly the average number of trades in a 15 minute interval. We also show error bars computed as specified by Gabaix *et al*. We cannot reject the null hypothesis of a linear relation between $E[r^2|V]$ and V with 95% confidence, even though we have a large amount of data, and we know by construction that β is quite different from 0.5.

Does it make a difference that we used a fixed number of transactions rather than a fixed time interval? To test this we

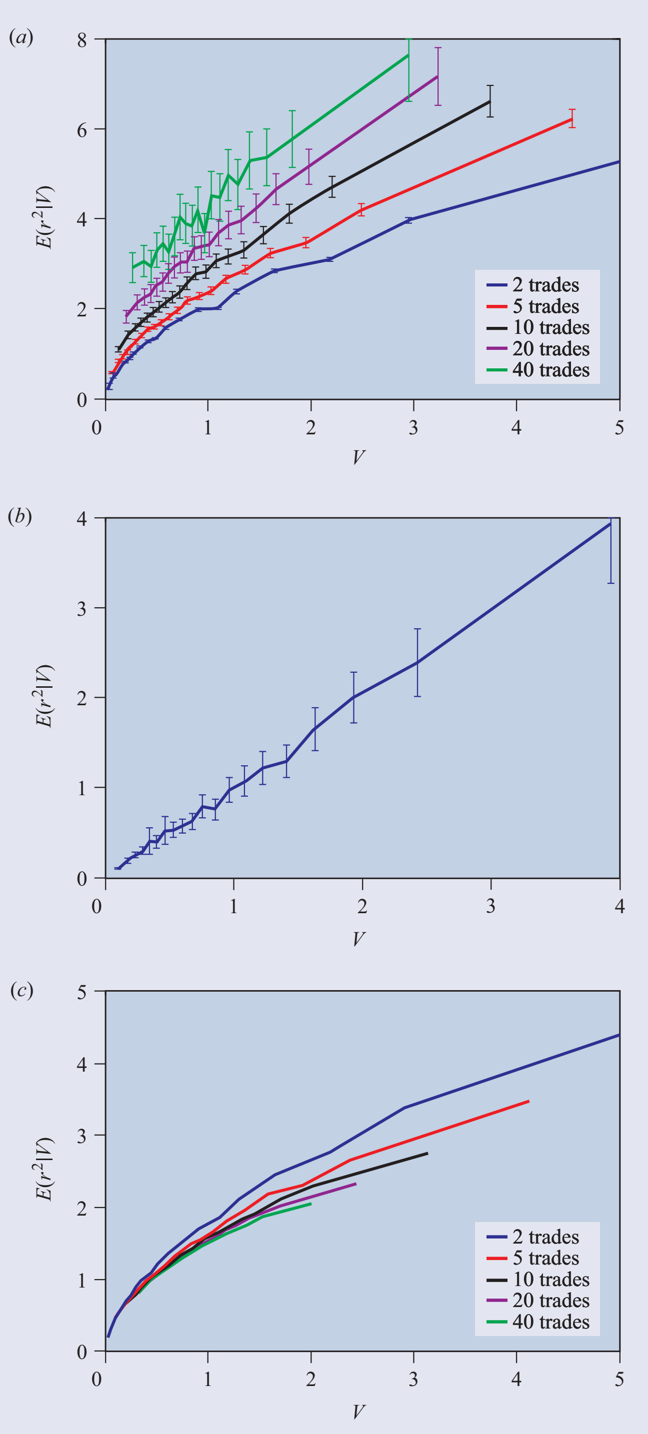

Figure 1. A demonstration that the statistical test of [1] fails due to the strong autocorrelations in real data. The expected value of the squared-price return, $E[r^2|V]$, is plotted against the total transaction size $V = \sum_{i=1}^{M} V_i$, where V_i is the size of transaction i. Each transaction causes a simulated market impact of the form $r_i = k\epsilon_i V_i^{\beta}$, to generate total return $r = \sum_{i=1}^{M} r_i$. Part (a) shows results for a fixed number of transactions, with M varying from 2 to 40; the curves are in ascending order of M; (b) is the same using a fixed time interval of 15 minutes, with variable M; and (c) is the same as (a) with the order of the transactions randomly shuffled. For (a) and (b) we see straight lines for large M, indicating that the test is passed, even though by construction the market impact does not follow the $r \sim V^{0.5}$ hypothesis, whereas for the shuffled data the test quite clearly shows us that the hypothesis is false.

repeat the procedure using a fixed time interval of 15 minutes. Figure 1(b) shows the result: an even clearer linear relation between $E[r^2|V]$ and V than before, so the test once again fails.

Why does this test not work? To gain some understanding of this, we repeat the same test but shuffle the order of the data, which breaks the correlation structure. As shown in figure 1(c), the result in this case is far from linear even when $M = 40$, and the test easily shows that the market impact does not follow a square-root law. Thus, we see that the problem lies in the autocorrelation structure of the real data.

In conclusion, our numerical simulations show that the linearity test of $E[r^2|V]$ lacks the power to test for a square-root market impact with data containing the correlation structure of real data. In fact, even a deterministic market impact like $r \sim V^{0.3}$ is consistent with the relation $E[r^2|V] = a + b\ V$ for a sufficiently large number of trades. Doing this for a fixed time interval rather than a fixed number of trades time makes this even more evident. Thus the Gabaix test provides no evidence that the average market impact follows a square-root law.

2. Placing error bars on the average market impact

In this section we present results for average market impact at the level of individual ticks. We show that it does not generally follow a square-root law, and that it varies from market to market, and in some cases from stock to stock, in a substantial and statistically significant way.

Realistic error bars for the average market impact are difficult to assess as volatility is a long-memory process [7,8]. That is, its time series has a slowly decaying power-law autocorrelation function that is asymptotically of the form $\tau^{-\kappa}$, with $\kappa < 1$ so that the integral is unbounded. This makes error analysis complicated, as data from the distant past have a strong effect on the present data. Because volatility is long memory, the price returns that fall in a given volume bin V_a, which are by definition all of the same sign, are also long memory. This means that the errors in measuring market impact are much larger than one would expect from intuition based on an IID hypothesis.

We analyse the market impact only for orders (or portions of orders) that result in immediate transactions[2]. Each transaction V_i generates a price return $r_i = \log p_a - \log p_b$, where p_b is the midpoint price quote just before the transaction and p_a is the midpoint price quote just after. We analyse buy and sell orders separately. The LSE has the advantage that the data set contains a record of orders, and so we can distinguish buy and sell orders unambiguously; for the NYSE data we use the trades and quotes to infer this using the Lee and Ready algorithm [9]. To estimate the average market impact we sort the events (V_i, r_i) with the same sign ϵ_i into bins based on V_i and plot the average value of V_i for each bin against the average value of r_i, as shown in figure 2. We choose the bins so that each bin has roughly the same number of points in it.

[2] *Note added in proof.* The LSE data that we use here contains only orders from the electronic (SETS) market. This has the advantage that the sign of orders (to buy or sell) is known with certainty. By 'transaction', throughout the paper we actually mean any component of an order that results in an immediate transaction. For the NYSE data the signs of orders must be inferred without complete certainty by comparison to the best quotes [9]; we lump together all transactions with the same time stamp and order code.

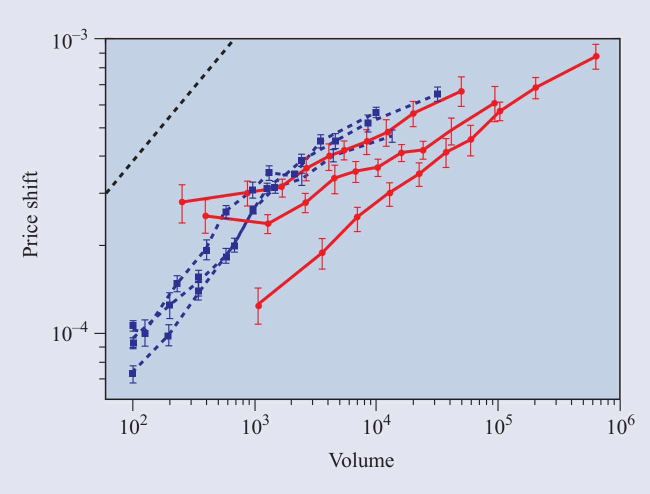

Figure 2. Market impact function for buy initiated trades of three stocks traded in the NYSE (dashed blue curve) and three stocks traded in the LSE (solid red curve). Trades of similar size V_i are binned together; on the horizontal axis we show the average volume of the trades in each bin, and on the vertical axis the average size of the logarithmic price change for the trades in that bin. In both cases comparison to the dashed black line in the corner, which has slope 0.5, makes it clear that the behaviour for large volume does not follow a law of the form $r_i \sim V_i^{0.5}$. Error bars are computed using the variance plot method [7] as described in the text.

To assign error bars for each bin we use the variance plot method [7]. For each bin we split the events into m subsamples with $n = K/m$ points, where K is the number of records in the bin. The subsamples are chosen to be blocks of values adjacent in time. For each subsample i we compute the mean $\mu_i^{(n)}$, $i = 1, \ldots, m$. Then we compute the standard deviation of the $\mu_i^{(n)}$ which we indicate as $\sigma^{(n)}$. By plotting $\sigma^{(n)}$ versus n in a log–log plot we compute the Hurst exponent H by fitting the data with a power-law function $\sigma^{(n)} = An^{H-1}$. We compute the error in the mean of the entire sample of K points by extrapolating the fitted function to the value $m = K$, i.e. $\sigma = \hat{A}\ K^{\hat{H}-1}$ where $\hat{A}$ and $\hat{H}$ are the ordinary least-squares estimates of the parameters A and H. Interestingly, for smaller values of V_i we find Hurst exponents substantially larger than 0.5, whereas for large values of V_i the Hurst exponents are much closer to 0.5. When $H > 0.5$ the error bars are typically much larger than standard errors[3].

In figure 2 we show empirical measurements of the average market impact for the NYSE and for the LSE. We consider three highly capitalized stocks for each exchange, Lloyds (LLOY), Shell (SHEL) and Vodafone (VOD) for the LSE and General Electric (GE), Procter & Gamble (PG) and AT&T (T) for the NYSE. For LSE stocks we consider the period May 2000–December 2002, while for NYSE stocks we consider the time period 1995–1996. The data for the NYSE are consistent with results reported earlier without error bars [3], while the LSE market impact data are new. The NYSE data clearly do not follow a power law across the whole range, consistent with earlier results in references [2,3]. While $\beta(V_i) \approx 0.5$ for small V_i, for larger V_i it appears that $\beta(V_i) < 0.2$. As shown in reference [3], this transition occurs for smaller values of V_i for stocks with lower capitalization. Thus, the assumption that $\beta = 0.5$ breaks down for high volumes, precisely where it is necessary in order for the theory of Gabaix *et al* to hold. For the LSE data the power-law assumption seems more justified across the whole range, but the exponent is too low; a least-squares fit gives $\beta \approx 0.26$. While we have not attempted to compute error bars for the regression, a visual comparison with the error bars of the individual bins makes it quite clear that $\beta = 0.5$ is inconsistent with either the LSE or NYSE data. It is also clear that the average market impact functions are qualitatively different for LSE and NYSE stocks, and that for NYSE stocks the functional form varies with market capitalization [3].

Even if we abandon the prediction that the average market impact is a square-root law, one can imagine that we could explain fluctuations in prices in terms of fluctuations in volume modulated by average market impact of the form $r_i = kV_i^{\beta}$. However, if this were true, for the NYSE the predicted exponent for price fluctuations would be $\alpha = \gamma/\beta \approx 1.5/0.25 = 6$, which is too large to agree with the data. (A typical value [10] is $\alpha \approx 3$.) To make matters worse, the power-law hypothesis for volume or market impact appears to fail in some other markets. In the PSE, Bouchaud *et al* [4] have suggested that the average market impact function[4] is of the form $\log V_i$, yielding $\beta \to 0$ in the limit as $V_i \to \infty$. For the LSE the power-law hypothesis for average market impact seems reasonable, but with an exponent significantly smaller than 0.5. Moreover, the volume is not power-law distributed, as discussed in the next section.

Note that we are making all the above statements for individual transactions, whereas many studies have been done based on aggregated data over a fixed time interval. Aggregating the data in time complicates the discussion, since the functional form of the market impact generally depends on the length of the time interval. Hence it is more meaningful to do the analysis based on individual transactions.

3. Volume distribution

The theory of Gabaix *et al* explains the power law of returns in terms of the power law of volume, so if volume does not have a power law, then, neither should returns. The existence of a power-law tail for volume seems to vary from exchange to exchange. For the NYSE we confirm the observation of power-law tails for volume reported earlier [11]. However in figure 3 we show the distribution of volumes for three stocks in the LSE. In order to compare different stocks we normalize the data by dividing by the sample mean for each stock.

All three stocks have strikingly similar volume distributions; this is true for the roughly 20 stocks that we have studied. There is no clear evidence for power-law scaling, even though the power-law scaling of the corresponding return distributions shown in figure 3(*b*) is clear. If one attempts to fit lines to the larger volume range of the curve (roughly 10^1–10^2), the exponent of the cumulative distribution corresponding to

[3] Since we choose the bins to have roughly the same number of points, the difference in Hurst exponent between bins with large and small V cannot be due to a difference in the mean interval between samples.

[4] For the NYSE the logarithmic form for average market impact is a reasonable approximation for small V_i, but breaks down for higher V_i.

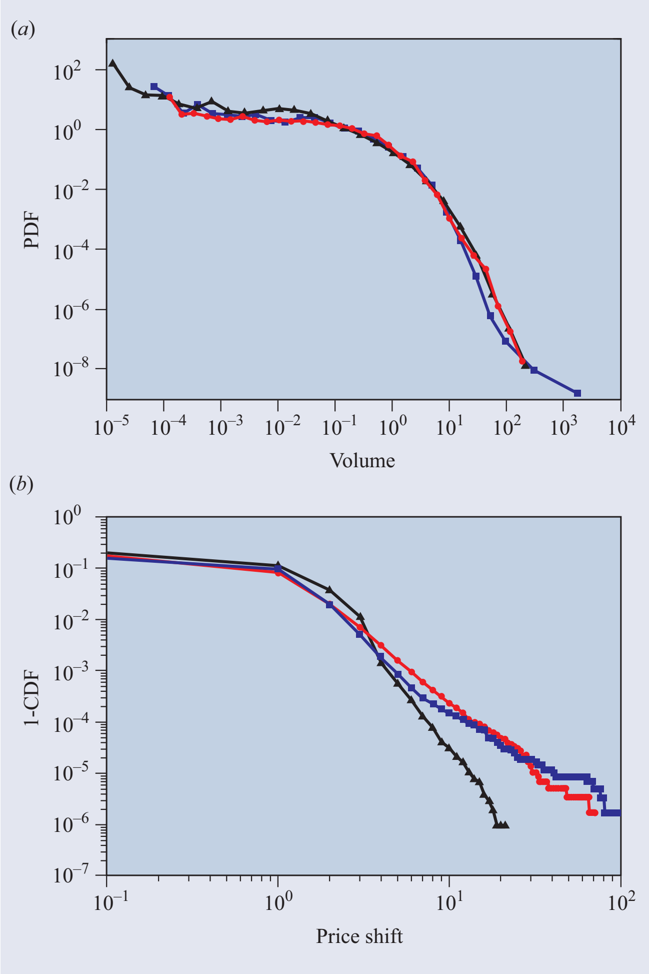

Figure 3. (a) The probability density of normalized volume for three typical high-volume stocks in the LSE, LLOY (red circles), SHEL (blue squares) and VOD (black triangles) in the period May 2000–December 2002. There are approximately 10^6 data points for each stock. (b) $1 - P(r)$, where $P(r)$ is the cumulative density function of returns induced by the same transactions in (a). For the normalized volume there is no clear evidence for power-law tails; in contrast for returns this is quite plausible. Furthermore, the volume distributions are essentially identical, whereas the return distribution for VOD decays more steeply than the others. Clearly one cannot explain the power-law scaling of returns based on that of volumes.

figure 3(a) is highly uncertain but it is at least 3, which, together with the measured values of β, would imply $\alpha \approx 3/0.3 \approx 10$. In contrast, the measured exponents for figure 3(b) are roughly 2.2, 2.5 and 4.3 for SHEL, LLOY and VOD respectively. It is noteworthy that VOD has a much larger α than the other stocks, even though it has essentially the same volume distribution and a similar volume distribution; if anything, from figure 2, its β is larger than that of the other stocks, which according to $\alpha = \gamma/\beta$ would imply a smaller α. This provides more evidence that the power-law tails of returns are not driven by those of volume.

4. Conclusion

We have shown that the conclusions of Gabaix *et al* [1] are invalid for three reasons. Firstly, their statistical analysis, in claiming the existence of a square-root law for average market impact, is invalid for the strong autocorrelations that are present in real data. Secondly, new measurements of the average market impact with proper error bars show that it does not follow a square-root law. Thirdly, for the LSE the distribution of volumes does not have a power-law tail, and there are substantial variations between the return distributions not reflected in variations in volume or average market impact. Thus, it seems quite clear that the distribution of large price fluctuations cannot be explained as a simple transformation of volume fluctuations.

This leaves open the question of what really causes the power-law tails of prices. We believe that the correct explanation lies in the extension of theories based on the stochastic properties of order placement and price formation [12–14], which naturally give rise to fluctuations in the response of prices to orders. Further work is clearly needed.

Note added in proof

The commentary [15] uncovered that trade size is distributed as a power law in the upstairs market, but not in the downstairs market, as conjectured in [6, 16]. Trades in the upstairs market (omitted from the electronic LSE data set used here) are arranged privately and then displayed only after they have already been made. Order splitting in the upstairs market is rare. Since this commentary was submitted 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 [16]. That is, if one matches up returns with the orders that generate them, the conditional distribution of large returns is essentially independent of order size. This was also seen for the NYSE and Island by Weber and Rosenow [17]. The idea that the tail of prices is driven by fluctuations in liquidity rather than fluctuations in the number of trades was implicitly suggested earlier by results of Plerou *et al* [18].

Acknowledgments

The authors thank the McKinsey Corporation, Credit Suisse First Boston, The McDonnel Foundation, Bob Maxfield and Bill Miller for their support, and Janos Kertesz, Moshe Levy, Rosario Mantegna and Ilija Zovko for valuable conversations.

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

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

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

[5] Bouchaud J-P, Gefen Y, Potters M and Wyart M 2003 Fluctuations and response in financial markets: the subtle nature of 'random' price changes *Preprint* cond-mat/0307332

[6] Lillo F and Farmer J D 2003 The long memory of the efficient market *Preprint* cond-mat/0311053

[7] Beran J 1994 *Statistics for Long-Memory Processes* (London: Chapman and Hall)
[8] Poon S-H and Granger C W J 2003 Forecasting volatility in financial markets: a review *J. Econ. Literature* **41** 478–539
[9] Lee C M C and Ready M J 1991 *J. Finance* **46** 733–46
[10] 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
[11] 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
[12] Daniels M, Farmer J D, Gillemot L, Iori G and Smith D E 2003 Quantitative model of price diffusion and market friction based on trading as a mechanistic random process *Phys. Rev. Lett.* **90** 108102
[13] Smith E, Farmer J D, Gillemot L and Krishnamurthy S 2003 Statistical theory of the continuous double auction *Quant. Finance* **3** 481–514
[14] Farmer J D, Patelli P and Zovko I I 2003 The predictive power of zero intelligence in financial markets *Preprint* cond-mat/0309233
[15] Plerou V, Gopikrishnan P, Gabaix X and Stanley H E 2004 On the origin of power-law fluctuations in stock prices *Quant. Finance* **4** C11-15
[16] Farmer J D, Gillemot L, Lillo F, Mike S and Sen A 2003 What really causes large price changes? *Preprint* cond-mat/0312703
[17] Weber P and Rosenow B 2004 Large stock price changes: volume or liquidity? *Preprint* cond-mat/0401123
[18] 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

J Doyne Farmer and Fabrizio Lillo are at the Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA.
Fabrizio Lillo is also at the Istituto Nazionale per la Fisica della Materia, Unità di Palermo, Italy.