arXiv:cond-mat/0307332 v1 14 Jul 2003

Fluctuations and response in financial markets: the subtle nature of 'random' price changes

Jean-Philippe Bouchaud[†,*], Yuval Gefen[×], Marc Potters[*], Matthieu Wyart[†]

July 14, 2003

[†] Commissariat à l'Energie Atomique, Orme des Merisiers
91191 Gif-sur-Yvette CEDEX, France

[*] Science & Finance, Capital Fund Management, 109-111 rue Victor Hugo
92 532 Levallois CEDEX, France

[×] Condensed Matter Physics Department, Weizmann Institute of Science,
76 100 Rehovot, Israel

July 14, 2003

Abstract

Using Trades and Quotes data from the Paris stock market, we show that the random walk nature of traded prices results from a very delicate interplay between two opposite tendencies: strongly correlated market orders that lead to super-diffusion (or persistence), and mean reverting limit orders that lead to sub-diffusion (or anti-persistence). We define and study a model where the price, at any instant, is the result of the impact of all past trades, mediated by a non constant 'propagator' in time that describes the response of the market to a single trade. Within this model, the market is shown to be, in a precise sense, at a critical point, where the price is purely diffusive and the average response function almost constant. We find empirically, and discuss theoretically, a fluctuation-response relation. We discuss the information content of each trade, and find that it is on average very small.

1 Introduction

The Efficient Market Hypothesis (EMH) posits that all available information is included in prices, which emerge at all times from the consensus between fully rational agents, that would otherwise immediately arbitrage away any deviation from the fair price [1, 2]. Price changes can then only be the result of unanticipated news and are by definition totally unpredictable. The price is at any instant of time the best predictor of future prices. One of the central predictions of EMH is thus that prices should be random walks in time which (to a good approximation) they indeed are. This was interpreted early on as a success of EMH. However, as pointed out by Schiller, the observed volatility of markets is far too high to be compatible with the idea of fully rational pricing [3]. The frantic activity observed in financial markets is another problem: on liquid stocks, there is typically one trade every 5 seconds, whereas the time lag between of relevant news is certainly much larger. More fundamentally, the assumption of rational, perfectly informed agents seems intuitively much too strong, and has been criticized by many [4, 5, 6]. Even the very concept of the fair price of a company appears to be somewhat dubious.

There is a model at the other extreme of the spectrum where prices also follow a pure random walk, but for a totally different reason. Assume that agents, instead of being fully rational, have zero intelligence and take random decisions to buy or to sell, but that their action is interpreted by all the others agents as potentially containing some information. Then, the mere fact of buying (or selling) typically leads to a change of the ask $a(t)$ (or bid $b(t)$) price and hence of a change of the midpoint $m(t) = [a(t)+b(t)]/2$. If the new midpoint is immediately adopted by all other market participants as the new reference price around which new orders are launched, then the midpoint will also follow a random walk (at least for sufficiently large times), even if trades are not motivated by any rational decision and devoid of meaningful information. This alternative, random trading model has been recently the object of intense scrutiny, in particular as a simplified approach to the statistics of order books [7, 8, 9, 10, 11, 12, 13, 14]. Since the order flow is a Poisson process, this assumption is quite convenient and leads to tractable analytical models [13, 15]. Perhaps surprisingly, many qualitative (and sometimes quantitative) properties of order books can be predicted using such an extreme postulate [11, 12, 13, 16].

Of course, reality should lie somewhere in the middle: clearly, the price cannot wander arbitrarily far from a reasonable value, and trades cannot all be random. The interesting question is to know which of the two pictures is closest to reality and can be taken as a faithful starting point around which improvements can be perturbatively added.

In this paper, we want to argue, based on a series of detailed empirical results obtained on trade by trade data, that the random walk nature of prices is in fact highly non trivial and results from a fine-tuned competition between two populations of traders, liquidity providers ('market-makers') on the one hand, and liquidity takers (sometimes called 'informed traders', but see the discussion in Section 4). For reasons that we explain in more details below, liquidity providers act such as to create anti-persistence (or mean reversion) in price changes that lead to a sub-diffusive behaviour of the price, whereas liquidity takers' action leads to persistence and super-diffusive behaviour. Both effects very precisely compensate and lead to an overall diffusive behaviour, at least to a first approximation. However, one can spot out the vestiges of this subtle compensation from the temporal structure of the market impact function (which measures how a given trade affects on average future prices).

The organization of this paper is as follows. We first present (Section 2) our empirical results on the statistics of trades, market impact and fluctuations. Then, we introduce in Section 3 a simple model that expresses the price as a linear superposition of the impact of each trade. We show that this model allows to rationalize our empirical findings, provided a stringent relation between the temporal autocorrelation of the sign of the trades (i.e. buyer initiated or seller initiated) and the temporal response to a single trade is satisfied. Finally, in Section 4, we give intuitive arguments that allow one to understand in financial terms the origin of this subtle balance between two opposite effects. We argue that in a very precise sense, the market is sitting on a critical point; this is similar to other complex systems such as the heart [17], that is driven by two conflicting systems (sympathetic and para-sympathetic), or certain human tasks, such as balancing of a long stick [18]. This suggests that any small deviation from this perfect balance may lead to strong instabilities, and that this near instability may be at the origin of the fat tails and volatility clustering observed in financial data (see e.g. [19, 20, 21, 22, 23, 24]).

2 Market impact and fluctuations

2.1 Presentation of the data and definitions

In this study, we have analyzed trades and quotes data from liquid French stocks in the years 2001 and 2002, although qualitatively similar results were also obtained on British stocks as well. The advantage of the French market, however, is that it is fully electronic whereas only part of the volume is traded electronically in the London stock exchange. We will illustrate our results mainly using the France-Telecom stock, which is one of the most actively traded stocks, for which statistics are particularly good.

There are two data files for each stock: one gives the list of all successive *quotes*, i.e. the best buy (bid, b) and sell (ask, a) prices, together with the available volume, and the time stamp accurate to the second. A quote can change either as a result of a trade, or because new limit orders appear, or else because some limit orders are canceled. The other data file is the list of all successive *trades*, with the traded price, traded volume and time stamp, again accurate to the second. Sometimes, several trades are recorded at the very same instant but at different prices: this corresponds to a market order of a size which exceeds the available volume at the bid (or at the ask), and hits limit orders deeper in the order book. In the following, we have grouped all these trades together as a single trade. This allows one to create chronological sequences of trades and quotes, such that between any two trades there is at least one quote.

The last quote before a given trade allows one to define the sign of each trade: if the traded price is above the last midpoint $m = (a + b)/2$, this means that the trade was triggered by a market order (or marketable limit order) to buy, and we will assign to that trade a variable $\varepsilon = +1$. If, one the other hand the traded price is below the last midpoint $m = (a + b)/2$, then $\varepsilon = -1$. With each trade is also associated a volume V, corresponding to the total number of shares exchanged.

Trades appear at random times, the statistics of which being itself non trivial (there are intra-day seasonalities and also clustering of the trades in time). We will not be interested in this aspect of the problem and always reason in terms of trade time, i.e. time advances by one unit every time a new trade (or a series of simultaneous trades) is recorded. We have also systematically discarded the first ten and the last ten minutes of trading in a given day, to remove any artifacts due to the opening and closing of the market. Many quantities of interest in the following are two-time observables, that is, compare two observables at (trade) time n and $n + \ell$. In order to avoid overnight effects, we have restricted our analysis to intra-day data, i.e. both n and $n + \ell$ belong to the same trading day. We have also assumed that our observables only depend on the time lag ℓ.

On the example of France-Telecom, on which we will focus mostly, there are on the order of 10 000 trades per day. For example, the total number of trades on France-Telecom during 2002 was close to $2.\,10^6$; this allows quite accurate statistical estimates of various quantities. The volume of each trade was found to be roughly log-normally distributed, with $\langle \ln V \rangle \simeq 5.5$ and a root mean square of $\Delta \ln V \sim 1.8$. The range of observed values of $\ln V$ is between 1 and 11.

2.2 Price fluctuation and diffusion

The simplest quantity to study is the average mean square fluctuation of the price between (trade) time n and $n + \ell$. Here, the price p_n is defined as the mid-point before the nth trade: $p_n \equiv m_{n^-}$. In this paper, we always consider detrended prices, such that the empirical drift is zero. We thus define $\mathcal{D}(\ell)$ as:

$$\mathcal{D}(\ell) = \left\langle (p_{n+\ell} - p_n)^2 \right\rangle . \tag{1}$$

As is well known, in the absence of any linear correlations between successive price changes, $\mathcal{D}(\ell)$ has a strictly diffusive behaviour, i.e.

$$\mathcal{D}(\ell) = D\ell, \tag{2}$$

where D is a constant. In the presence of short-ranged correlations, one expects deviations from this behaviour at short times. Quite surprisingly, however, on

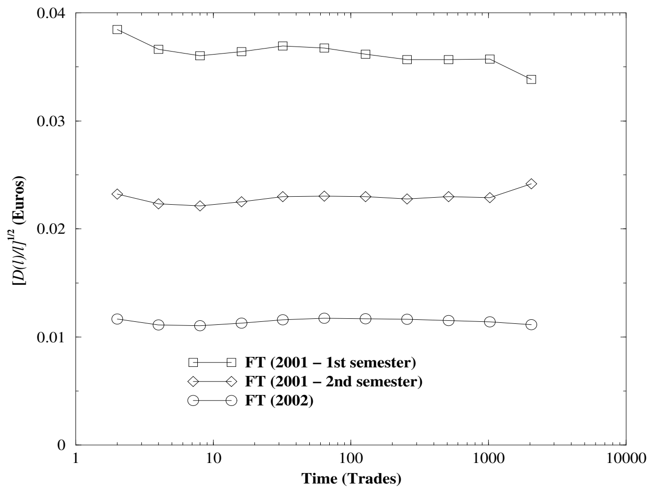

Figure 1: Plot of $\sqrt{\mathcal{D}(\ell)/\ell}$ as a function of ℓ for France-Telecom, during three different periods. The variation of $\mathcal{D}(\ell)/\ell$ with ℓ is very small, in particular in the small tick (0.01 Euros) period (July 2001 – December 2002). For the large tick size period (0.05 Euros; January 2001 – June 2001), there is a systematic downward trend: see also Fig. 2.

liquid stocks with relatively small tick sizes such as France-Telecom (FT), one finds a remarkably linear behaviour for $\mathcal{D}(\ell)$, even for small ℓ. In fact, in order to emphasize the differences from a strictly diffusive behaviour, we have studied the quantity $\sqrt{\mathcal{D}(\ell)/\ell}$ which has the dimension of Euros. We show this quantity in Fig. 1 for FT, averaged over three different periods: first semester of 2001 (where the tick size was 0.05 Euros), second semester of 2001, and the whole of 2002 (where the tick size was 0.01 Euros). One sees that $\mathcal{D}(\ell)/\ell$ is indeed nearly constant, with a small 'oscillation' on which we will comment later. Similar plots can be observed for other stocks (see Fig. 2). We have noted that for stocks with larger ticks, a slow decrease of $\mathcal{D}(\ell)/\ell$ is observed, corresponding to a slight anti-persistence (or sub-diffusion) effect.

The conclusion is that the random walk (diffusive) behaviour of stock prices appears even at the trade by trade level, with a diffusion constant D which is of the order of the typical bid-ask squared. From Fig. 1, one indeed sees that $\sqrt{\mathcal{D}(1)} \sim 0.01$ Euros, which is precisely the tick size, and FT has a typical bid-ask spread equal to one or two ticks. This coincidence is interesting. It might suggests that price changes are to a large extent induced by the trading activity itself, independently of real news (unless of course if the news flow is itself on the scale of seconds and that each news item has an impact on the price that is commensurate to the bid-ask spread). Much stronger arguments in favor of this point, based on estimates of the information content of each trade, will be given

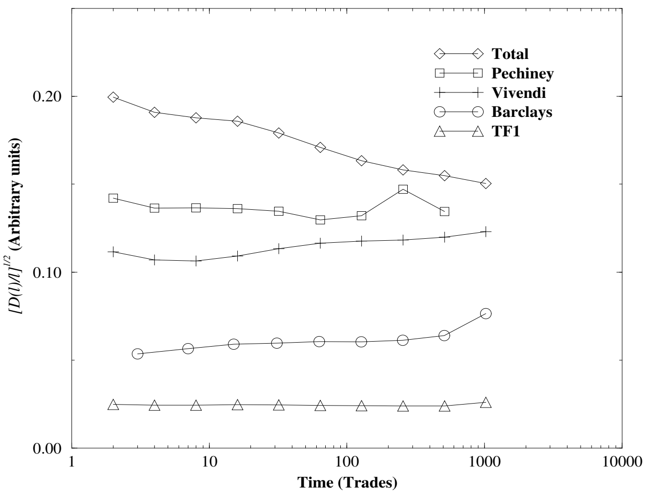

Figure 2: Plot of $\sqrt{\mathcal{D}(\ell)/\ell}$ as a function of ℓ for other stocks during the year 2002, except Barclays (May-June 2002). The y-axis has been rescaled arbitrarily for clarity. We note that stocks with larger tick size tend to reveal a stronger mean-reverting effect.

below. This conclusion seems to imply that the price may, on the long run, wander arbitrarily far from the fundamental price, which would be absurd. However, even if one assumes that the fundamental price is independent of time, a typical 3% noise induced daily volatility would lead to a factor two difference between the traded price and the fundamental price only after a few years [25]. Since the fundamental price of a company is probably difficult to determine better than within a factor two, say (see e.g. [5, 26]), one only expects fundamental effects to be relevant on very long time scales (as indeed suggested by the empirical results of de Bondt and Thaler [27]), but that these are totally negligible on short (intra-day) time scales of interest here. We will in fact see below that the reference price that market participants seem to have in mind is in fact a short time average of the past price itself, rather than any fundamental price.

2.3 Response function and market impact

In order to better understand the impact of trading on price changes, one can study the following *response function* $\mathcal{R}(\ell)$, defined as:

$$\mathcal{R}(\ell) = \langle (p_{n+\ell} - p_n) \cdot \varepsilon_n \rangle \,, \tag{3}$$

where ε_n is the sign of the n-th trade, introduced in Section 2.1. The quantity $\mathcal{R}(\ell)$ measures how much, on average, the price moves up conditioned to a buy order at time 0 (or a sell order moves the price down) a time ℓ later. As will be clear below,this quantity is however not the market response to a single trade, a quantity that will later be denoted by G_0. A more detailed object can in fact be defined by conditioning the average to a certain volume V of the n-th trade:

$$\mathcal{R}(\ell, V) = \langle (p_{n+\ell} - p_n) \cdot \varepsilon_n \rangle|_{V_n=V} \,. \tag{4}$$

Previous empirical studies have mostly focused on the volume dependence of the short time limit ($\ell \sim 1$) of $\mathcal{R}(\ell, V)$, and established that this function is strongly concave as a function of the volume [28, 29]. In [30], a thorough analysis of U.S. stocks have lead the authors to propose a power-law dependence for $\mathcal{R}(\ell = 1, V) \propto V^\alpha$, with an exponent $\alpha \simeq 0.4$ for small volumes, and a smaller value still ($\alpha \simeq 0.2$) for larger volumes. In a previous publication [31], some of us have proposed that this dependence might in fact be logarithmic (see also [29]): $\mathcal{R}(\ell = 1, V) = R_1 \ln V$ (where R_1 is a stock dependent constant), a law that seems to satisfactorily account for all the data that we have analyzed. The empirical determination of the temporal structure of $\mathcal{R}(\ell, V)$ has been much less investigated (although one can find in [29] somewhat related results on a coarse-grained version of $\mathcal{R}(\ell, V)$). Preliminary empirical results, published in [31], reported that $\mathcal{R}(\ell, V)$ could be written in a factorized form:

$$\mathcal{R}(\ell, V) \approx \mathcal{R}(\ell) f(V); \qquad f(V) \propto \ln V, \tag{5}$$

(this factorized form was actually first suggested on theoretical grounds in [11]), where $\mathcal{R}(\ell)$ is a slowly varying function that initially increases up to $\ell \sim 100-1000$ and then is seen to decrease back, with a rather small overall range of variation. The initial increase of $R(\ell)$ has also recently been noticed by Farmer and Lillo [16]. Here, we provide much better data that supports both the above assertions. We show for example in Fig. 3 the temporal structure of $\mathcal{R}(\ell)$ for France Telecom, for different periods. Note that $\mathcal{R}(\ell)$ increases by a factor ~ 2 between $\ell = 1$ and $\ell = \ell^* \approx 1000$, before decreasing back. Similar results have been obtained for many different stocks as well: Fig. 4 shows a small selection of other stocks, where the non monotonous behaviour of $\mathcal{R}(\ell)$ is shown. However, in some cases (such as Pechiney), the maximum is not observed. One possible reason is that the number of daily trades is in this case much smaller (~ 1000), and that ℓ^* is beyond the maximum intra-day time lag.

The existence of a time scale ℓ^* beyond which $\mathcal{R}(\ell)$ decreases is thus both statistically significant, and to a large degree independent of the considered stock. On the other hand, the amplitude of the change of $\mathcal{R}(\ell)$ seems to be stock dependent. As will be clear later, the slowly varying nature of $\mathcal{R}(\ell)$ and the fact that this quantity reaches a maximum are non trivial results that will require a specific interpretation.

Turning now to the factorization property of $\mathcal{R}(\ell, V)$, Eq. (5), we illustrate its validity in Fig. 5, where $\mathcal{R}(\ell, V)/f(V)$ is plotted as a function of ℓ for different values of V. The function $f(V)$ was chosen for best visual rescaling, and is found to be close to $f(V) = \ln V$, as expected. Note that for the smallest volume

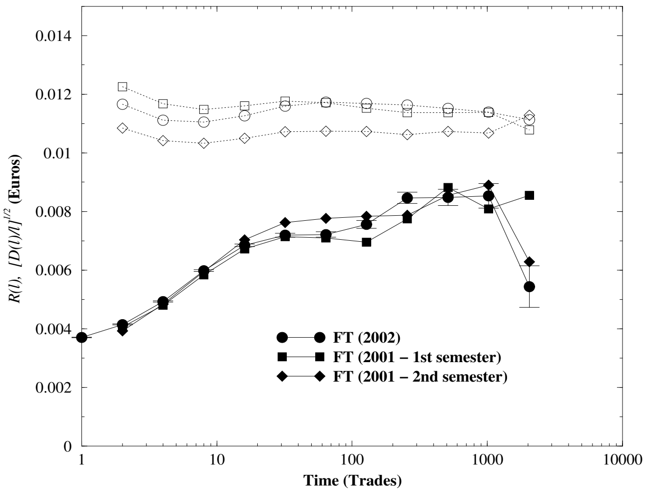

Figure 3: Average response function $\mathcal{R}(\ell)$ for FT, during three different periods (black symbols). We have given error bars for the 2002 data. For the 2001 data, the $y-$axis has been rescaled to best collapse onto the 2002 data. Using the same rescaling factor, we have also shown the data of Fig. 1. The fact that this rescaling works approximately will be dwelled further in Section 2.4 below.

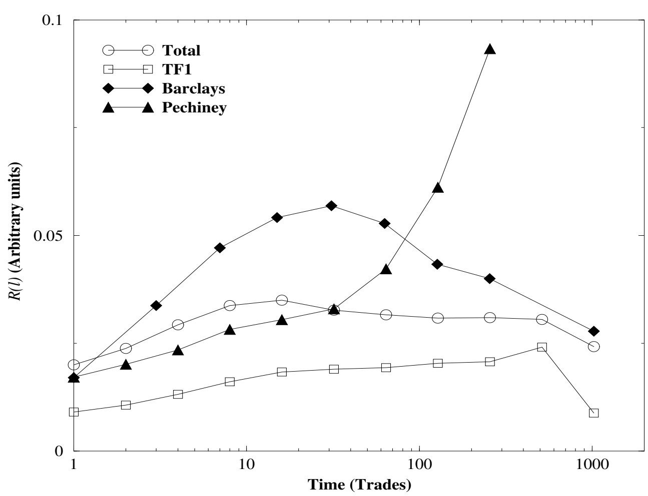

Figure 4: Average response function $\mathcal{R}(\ell)$ for a restricted selection of stocks, during 2002.

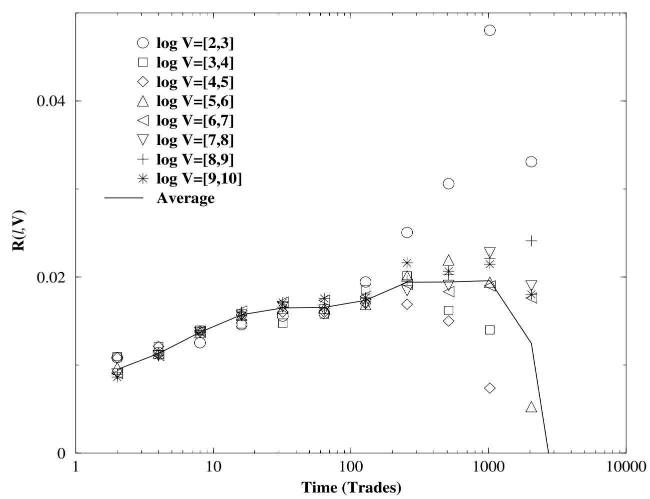

Figure 5: Average response function $\mathcal{R}(\ell, V)$, conditioned to a certain volume V, as a function of ℓ. Data for different V's have been divided by $f(V) \propto \ln V$ such as to obtain good data collapse. The thick line corresponds to $\mathcal{R}(\ell)$ (unscaled).

(open circles), the long time behaviour of $\mathcal{R}(\ell, V)$ seems to be different, which is probably due to the fact that small volumes are in fact more likely to be large volumes chopped up into small pieces.

One has to keep in mind that the response function $\mathcal{R}(\ell)$ captures a small systematic effect that relates the average price change to the sign of a trade. However, the fluctuations around this small signal are large, and increase with ℓ. A way to see this is to introduce the random variable $u_\ell = (p_{n+\ell} - p_n).\varepsilon_n$. By definition, $\mathcal{R}(\ell)$ is the average of u_ℓ, and $\mathcal{D}(\ell)$ is the average of u_ℓ^2. Since $\mathcal{R}(\ell)$ is roughly constant whereas $\mathcal{D}(\ell)$ grows linearly with ℓ, one sees that the impact of a given trade (as measured by $\mathcal{R}(\ell)$) rapidly becomes lost in the fluctuations. In Fig. 6, we show the whole empirical distribution $P(u_\ell)$ of u_ℓ for $\ell = 128$. This distribution is found to be only slightly skewed in the direction of positive u_ℓ. In fact, if one considers the shifted variable $u_\ell - \nu$, where $\nu = 0.02$ Euros, the distribution becomes nearly symmetric. Note that 0.02 is roughly equal to the typical bid-ask spread and can therefore be seen as the cost of a market order. The Efficient Market picture would suggest that the non zero value of $\langle u_\ell \rangle$ is mostly due to a small fraction of trades that move the price a lot as a result of some real information, while most 'noise' induced trades would not change the price at all. In this case, the positive tail of the distribution $P(u_\ell)$ (corresponding to informed trades) should be much fatter than the negative tail. This asymmetry could in fact be taken as an objective measure of the fraction of informed trades. The nearly symmetric shape of $P(u_\ell)$ shown in Fig. 6 means that one can hardly detect the statistical presence of informed trades that correctly anticipate the sign of the price change on a short term basis. Note that $P(u_\ell)$ gives an equal

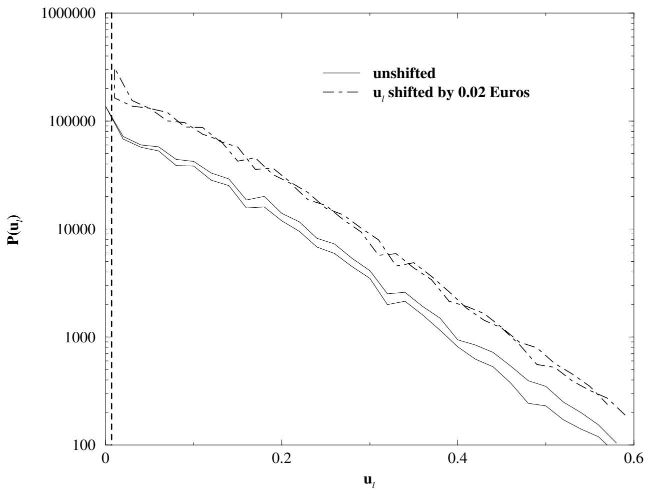

Figure 6: Probability distribution $P(u_\ell)$ of the quantity $u_\ell = (p_{n+\ell} - p_n).\varepsilon_n$, for $\ell = 128$. The data is again FT during 2002. The negative part of the distribution has been folded back to positive u_ℓ in order to highlight the small positive skew of the distribution (which is seen to increase slightly with $|u_\ell|$). The average value $\mathcal{R}(\ell) = \langle u_\ell \rangle$ is shown as the vertical dashed line. The dashed-dotted line corresponds to the distribution of u_ℓ shifted by 0.02 Euros. This curve has been shifted vertically for clarity.

weight to all trades, independently of their volume, but considering the volume weighted distribution leads to the same conclusion.

The conclusion of this section are thus that (a) large volumes impact prices much less (in relative terms) than smaller volumes and (b) the average impact of a given trade (as measured by $\mathcal{R}(\ell)$) increases with time up to a certain time scale ℓ^* beyond which it decreases.

2.4 A Fluctuation-Response relation

In the study of Brownian particles, a very important result that dates back to Einstein relates the diffusion coefficient D to the response of the particle to an external force. That a similar relation might also hold in financial markets was first suggested by Rosenow [32], and substantiated there by some empirical results. We have performed a similar analysis within the framework and notations of the present paper: for any given trading day, one can compute the average local diffusion constant $\mathcal{D}(\ell)$ for a given time scale, say $\ell = 128$, and the average local price response $\mathcal{R}(\ell)$ over the same time scale. The analogue of Rosenow's result [32] (which was motivated by a Langevin equation for price variations (see [33])), is a linear relation between $\mathcal{R}^2(\ell)$ and $\mathcal{D}(\ell)$, which we illustrate in Fig. 7 for FT, for two different periods (first semester of 2001, and 2002). A similar result can

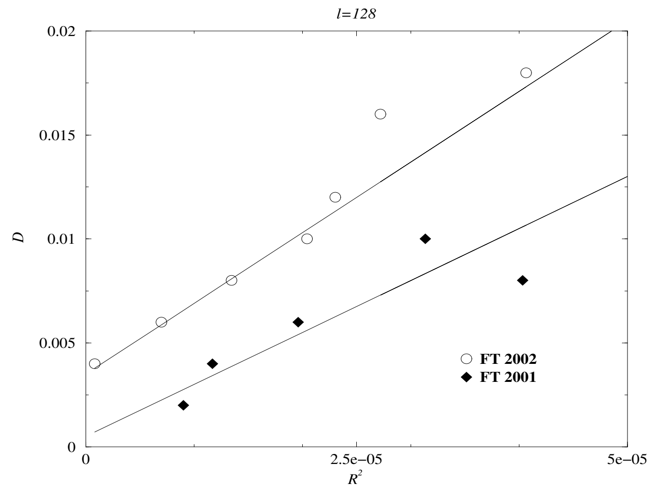

Figure 7: Average diffusion constant $\mathcal{D}(\ell)/\ell$, computed for $\ell = 128$, and conditioned to a certain value of $\mathcal{R}^2(\ell)$, also computed for $\ell = 128$ (FT). The open symbols correspond to 2002, whereas the black symbols are computed using the first semester of 2001, where the tick size was 5 times larger. Correspondingly, the x-axis was rescaled by a factor 25 and the y-axis by a factor five for this data set.

also be read from Fig. 3. As will be clear in the following, such a relation will appear naturally within the simple model that we introduce in Section 3.

2.5 Long term correlation of trade signs

All the above results are compatible with a 'zero intelligence' picture of financial markets, where each trade is random in sign and shifts the price permanently, because all other participants update their evaluation of the stock price as a function of the last trade. As shown in [8, 9, 11, 12, 13, 14], a model of the order book based on a purely random order flow indeed allows one to go quite far in the quantitative understanding of financial markets. In particular, the concave shape of the impact as a function of the volume can be understood as an order book effect, where the average size of the queue increases with depth.

This model of a totally random stock market is however qualitatively incorrect for the following reason. Although, as mentioned above, the statistics of price changes reveals very little temporal correlations, the correlation function of the sign ε_n of the trades, on the other hand, reveals very slowly decaying correlations. More precisely, one can consider the following correlation function: [1]

$$\mathcal{C}_0(\ell) = \langle \varepsilon_{n+\ell} \varepsilon_n \rangle \tag{6}$$

[1] Assuming, as is the case empirically, that $\langle \varepsilon_n \rangle \simeq 0$.

If trades were random, one should observe that $\mathcal{C}(\ell)$ decays to zero beyond a few trades. Surprisingly, this is not what happens: on the contrary, $\mathcal{C}(\ell)$ is strong and decays very slowly toward zero, as an inverse power-law of ℓ (see Fig. 8):

$$\mathcal{C}_0(\ell) \simeq \frac{C_0}{\ell^\gamma}, \qquad (\ell \geq 1). \tag{7}$$

The value of γ seems to be somewhat stock dependent. For example, for FT, one finds $\gamma \approx 1/5$, whereas for Total $\gamma \approx 2/3$. Vodafone is characterized by a somewhat larger value of $\gamma \approx 1/2$ [16]. In any case, the value of γ is found to be smaller than one, which is very important because the integral of $\mathcal{C}_0(\ell)$ is then *divergent*. Now, as will be shown more precisely in the next section, the integral of $\mathcal{C}_0(\ell)$ can intuitively be thought of as the effective number N_e of correlated successive trades. Hence, out of – say – 1000 trades, one should group together

$$N_e \simeq 1 + \sum_{\ell=1}^{1000} \mathcal{C}_0(\ell) \approx 1 + \frac{C_0}{1-\gamma} 1000^{1-\gamma} \tag{8}$$

'coherent' trades. For FT, $\gamma \approx 1/5$ and $C_0 \approx 0.2$, which means that the effect of one trade should be amplified, through the correlations, by a factor $N_e \approx 50$! In other words, both the response function $\mathcal{R}$ and the diffusion constant should increase by a factor 50 between $\ell = 1$ and $\ell = 1000$, in stark contrast with the observed empirical data. This is the main paradox that one should try to elucidate: how can one reconcile the strong, slowly decaying correlations in the sign of the trades with the nearly diffusive nature of the price fluctuations, and the nearly structureless response function?

Before presenting a mathematical transcription of the above question and proposing a possible resolution, let us comment on two related correlation functions that will naturally appear in the following, namely:

$$\mathcal{C}_1(\ell) = \langle \varepsilon_{n+\ell}\, \varepsilon_n \ln V_n \rangle, \tag{9}$$

and

$$\mathcal{C}_2(\ell) = \langle \varepsilon_{n+\ell} \ln V_{n+\ell}\, \varepsilon_n \ln V_n \rangle. \tag{10}$$

We have found empirically that these two 'mixed' correlation functions are proportional to $\mathcal{C}_0(\ell)$ (see Fig 8):

$$\mathcal{C}_1(\ell) \approx \langle \ln V \rangle \mathcal{C}_0(\ell); \qquad \mathcal{C}_2(\ell) \approx \langle \ln V \rangle^2 \mathcal{C}_0(\ell). \tag{11}$$

There are however small systematic deviations, which indicate that small volumes contribute more to the long range correlations that larger volumes.

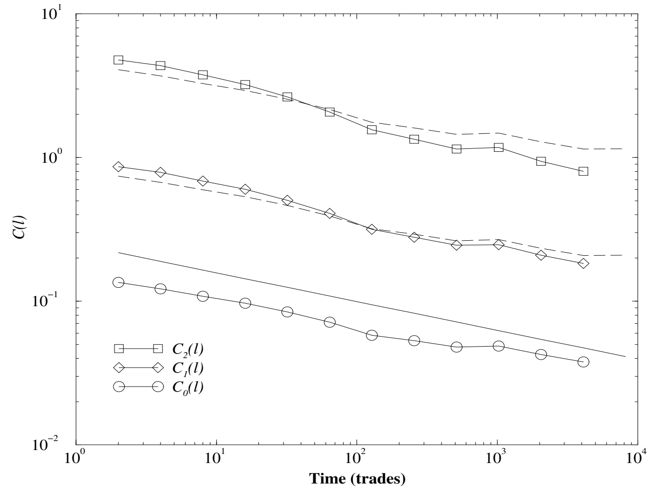

Figure 8: Volume weighted sign autocorrelation functions as a function of time lag: $\mathcal{C}_0$, $\mathcal{C}_1$, $\mathcal{C}_2$ (see text for definitions). The straight line corresponds to C_0/ℓ^γ with $\gamma = 1/5$. The dotted lines correspond to the simple approximation given by Eq. (11).

3 A micro-model of price fluctuations

In order to understand the above results, we will postulate the following trade superposition model, where the price at time n is written as a sum over all past trades, of the impact of one given trade propagated up to time n:

$$p_n = \sum_{n'<n} G_0(n-n')\varepsilon_{n'} \ln V_{n'} + \sum_{n'<n} \eta_{n'}, \tag{12}$$

where $G_0(.)$ is the 'bare' impact function (or propagator) of a single trade, that we assume to be a fixed, non random function that only depends on time differences. The η_n are also random variables, assumed to be independent from the ε_n and model all sources of price changes not described by the direct impact of the trades: the bid-ask can change as the result of some news, or of some order flow, in the absence of any trades.

The bare impact function $G_0(\ell)$ represents by definition the average impact of a single trade after ℓ trades. It could be in principle measured empirically by launching on the market a sequence of real trades of totally random signs, and averaging the impact over this sample of trades (a rather costly experiment!).[2] As will be clear below, the difference between the quantity $\mathcal{R}(\ell)$ introduced in the previous Section and $G_0(\ell)$ in fact comes from the strong autocorrelation of the sign of the trades. In order to understand the temporal structure of $G_0(\ell)$,

[2]However, following this procedure might induce 'copy-cat' trades and still lead to a difference between $\mathcal{R}$ and G_0

note that a single trade first impacts the midpoint by changing the bid (or the ask). But then the subsequent limit order flow due to that particular trade might either center on average around the new midpoint (in which case $G_0(\ell)$ would be constant), or, as we will argue below, tend to mean revert toward the previous midpoint (in which case $G_0(\ell)$ decays with ℓ). The asymptotic behaviour of the bare impact function in fact reveals the average informational content of a single trade: if $G_0(\ell \gg 1)/G_0(1)$ is close to unity, this informational content is large; if this ratio is small, so is the informational content, since the impact of the trade was only temporary, and was not followed by a long term change of the price.

Using this representation, the price increment between an arbitrarily chosen initial time 0 and time ℓ is:

$$p_\ell - p_0 = \sum_{0 \le n < \ell} G_0(\ell - n)\varepsilon_n \ln V_n + \sum_{n<0} \left[G_0(\ell - n) - G_0(-n)\right] \varepsilon_n \ln V_n + \sum_{0 \le n < \ell} \eta_n. \tag{13}$$

If the signs ε_n were independent random variables, both the response function and the diffusion would be very easy to compute. For example, one would have:

$$\mathcal{R}(\ell) = \langle \ln V \rangle G_0(\ell), \tag{14}$$

i.e. the observed impact function and the bare response function would be proportional. Similarly, one would have:

$$\mathcal{D}(\ell) = \langle \ln^2 V \rangle \left(\sum_{0<n\le\ell} G_0^2(n) + \sum_{n>0} \left[G_0(\ell + n) - G_0(n)\right]^2 \right) + D_\eta \ell, \tag{15}$$

where D_η is the variance of the η's. In the simplest case of a constant bare impact function, $G_0(\ell) = G_0$ for all $\ell > 0$, one then finds a pure diffusive behaviour, as expected:

$$\mathcal{D}(\ell) = \ell \left[\langle \ln^2 V \rangle G_0^2 + D_\eta \right]. \tag{16}$$

This result (no correlations between the ε's and a constant bare impact function) corresponds to the simplest possible zero intelligence market. However, we have seen that in fact the ε's have long range correlations. In this case, the average response function reads:

$$\mathcal{R}(\ell) = \langle \ln V \rangle G_0(\ell) + \sum_{0<n<\ell} G_0(\ell - n)\mathcal{C}_1(n) + \sum_{n>0} \left[G_0(\ell + n) - G_0(n)\right] \mathcal{C}_1(n). \tag{17}$$

Note in passing that our trade superposition model, Eq. (12), together with Eq. (11) leads to the factorization property mentioned above:

$$\mathcal{R}(\ell, V) = \frac{\ln V}{\langle \ln V \rangle} \mathcal{R}(\ell). \tag{18}$$

Now, one sees more formally the paradox discussed in the previous Section: assuming that the impact of each trade is permanent, i.e. $G_0(\ell) = G_0$, leads to:

$$\mathcal{R}(\ell) = \langle \ln V \rangle G_0 \left[1 + \sum_{0<n<\ell} \mathcal{C}_0(n) \right]. \tag{19}$$

If $\mathcal{C}_0(n)$ decays as a power-law with an exponent $\gamma < 1$, then the average impact $\mathcal{R}(\ell)$ should *grow* like $\ell^{1-\gamma}$, and therefore be amplified by a very large factor as ℓ increases. The only way out of this conundrum is (within the proposed model) that the bare impact function $G_0(\ell)$ itself should decay with time, in such a way to offset the amplification effect due to the trade correlations.

In order to get some guidance, let us now look at the general formula for the diffusion. After a few lines of calculations, one finds:

$$\begin{aligned} \mathcal{D}(\ell) &= \langle \ln^2 V \rangle \left[\sum_{0 \le n < \ell} G_0^2(\ell - n) + \sum_{n>0} \left[G_0(\ell+n) - G_0(n) \right]^2 \right] \\ &+ 2\Delta(\ell) + D_\eta \ell, \end{aligned} \tag{20}$$

where $\Delta(\ell)$ is the correlation induced contribution:

$$\begin{aligned} \Delta(\ell) &= \sum_{0 \le n < n' < \ell} G_0(\ell - n) G_0(\ell - n') \mathcal{C}_2(n' - n) \\ &+ \sum_{0<n<n'} \left[G_0(\ell+n) - G_0(n) \right] \left[G_0(\ell+n') - G_0(n') \right] \mathcal{C}_2(n' - n) \\ &+ \sum_{0 \le n < \ell} \sum_{n'>0} G_0(\ell - n) \left[G_0(\ell+n') - G_0(n') \right] \mathcal{C}_2(n' + n). \end{aligned} \tag{21}$$

Note that $\mathcal{D}(\ell)$ can alternatively be written in a more compact way as:

$$\mathcal{D}(\ell) = 4 \sum_{n \ge 0} \chi(n) \left[\mathcal{C}_2(n) - \mathcal{C}_2(n+\ell) \right] \tag{22}$$

with $\chi(\ell) \equiv \sum_{n \ge 0} G_0(n) G_0(\ell+n)$, provided this quantity exists.

The constraint from empirical data is that this expression must be approximately linear in ℓ. If we make the simple ansatz that the bare impact function $G_0(\ell)$ also decays as a power-law:

$$G_0(\ell) = \frac{R_0}{(\ell_0 + \ell)^\beta} \qquad (\ell \ge 1) \tag{23}$$

then one can estimate $\mathcal{D}(\ell)$ in the large ℓ limit. When $\gamma < 1$, one again finds that the correlation induced term $\Delta(\ell)$ is dominant, and all three terms scale a $\ell^{2-2\beta-\gamma}$, provided $\beta < 1$. In other words, the Hurst exponent of price changes is given by $2H = 2 - 2\beta - \gamma$. Therefore, the condition that the fluctuations are purely diffusive ($H = 1/2$) imposes a relation between the decay of the sign autocorrelation γ and the decay of the bare impact function β that reads:

$$2\beta + \gamma = 1 \longrightarrow \beta_c = \frac{1-\gamma}{2} \tag{24}$$

For $\beta > \beta_c$, the price is *sub-diffusive* ($H < 1/2$), which means that price changes show anti-persistence; while for $\beta < \beta_c$, the price is *super-diffusive* ($H > 1/2$), i.e. price changes are persistent. For FT, $\gamma \approx 1/5$ and therefore $\beta_c \approx 2/5$.

This is however still an apparent contradiction, for if one goes back to the response function given by Eq. (17), one finds that if $\beta + \gamma < 1$ (which is indeed the case for $\beta = \beta_c$), the dominant contribution to $\mathcal{R}(\ell)$ should behave as $\ell^{1-\beta-\gamma}$ and thus *grow* with ℓ. For example, for $\gamma \approx 1/5$ and $\beta \approx 2/5$, one should find that $\mathcal{R}(\ell) \propto \ell^{2/5}$, which is incompatible with the empirical data of Figs. 3 and 4. But if one now computes the numerical prefactor of this power law, one finds, for large ℓ:

$$\mathcal{R}(\ell) \simeq \langle \ln V \rangle R_0 C_0 \frac{\Gamma(1-\gamma)}{\Gamma(\beta)\Gamma(2-\beta-\gamma)} \left[\frac{\pi}{\sin \pi\beta} - \frac{\pi}{\sin \pi(1-\beta-\gamma)} \right] \ell^{1-\beta-\gamma}. \tag{25}$$

Therefore, when $2\beta + \gamma = 1$, the prefactor is exactly zero, and leads to the possibility of a nearly constant impact function! For faster decaying impact functions (larger β's), this prefactor is negative, whereas for more slowly decaying impact functions this prefactor is positive.[3] Interestingly, even if the bare response function $G_0(\ell)$ is positive for all ℓ, the average response $\mathcal{R}(\ell)$ can become negative for large enough β's, as a consequence of the correlations between trades.

Since the dominant term is zero for the 'critical' case $2\beta+\gamma = 1$, and since we are interested in the whole function $\mathcal{R}(\ell)$ (including the small ℓ regime), we have computed $\mathcal{R}(\ell)$ numerically, by performing the discrete sum Eq. (17) exactly. The results are shown in Fig. 9, for $\gamma = 1/5$ and three different values of β in the immediate vicinity of $\beta_c = 2/5$. We have fixed C_0 to its empirical value $C_0 = 0.2$ (see Fig. 8) and chosen $\ell_0 = 20$ as a free parameter. The overall scaling parameter R_0 is then adjusted to $R_0 = 2.8\,10^{-3}$ Euros. The results are compared with the empirical data for FT, showing that one can indeed satisfactorily reproduce, when $\beta \approx \beta_c$, a weakly increasing impact function that reaches a maximum and then decays. One also sees, from Fig. 9, that the relation between β and γ must be quite accurately satisfied, otherwise the response function shows a distinct upward trend (for $\beta < \beta_c$) or a downward trend ($\beta > \beta_c$).[4] In fact, we have tried other simple forms for $G_0(\ell)$, such as a simple exponential decay toward a possibly non zero asymptotic value, but this leads to unacceptable shapes for $\mathcal{R}(\ell)$.

[3]Note that although this prefactor increases (in absolute value) with β for $\beta > \beta_c$, the power of ℓ decreases, which means that for large ℓ the amplitude of $\mathcal{R}(\ell)$ decreases with β, as intuitively expected.

[4]This might actually explain the different behaviour of Pechiney seen in Fig. 4.

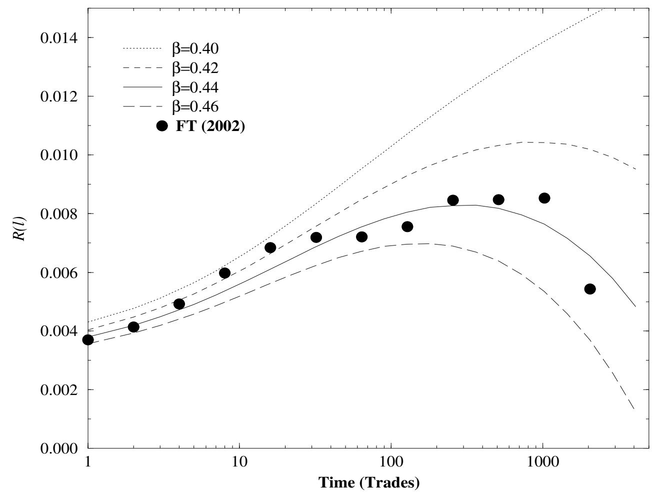

Figure 9: Theoretical impact function $\mathcal{R}(\ell)$, from Eq. (17), and for different values of β close to $\beta_c = 2/5$. The shape of the empirical response function can be quite accurately reproduced using $\beta = 0.44$. The only remaining free parameter in $\ell_0 = 20$.

As we showed above, the reason for the fine tuning of β is the requirement that price changes are almost diffusive. We can therefore also compute $\mathcal{D}(\ell)$ for all values of ℓ using the very same values of γ, β, $C_0\ell_0$ and R_0. Now, in order to fit the data one has two extra free parameters: one is D_η, and the other comes about because the mid-point can change without any trade. One should thus add to $\mathcal{D}(\ell)$ an ℓ-independent 'error' term D_0 that survives in the $\ell = 0$ limit, and is associated to bid-ask fluctuations. With these two extra parameters, one can reproduce very accurately the empirical deviations of $\mathcal{D}(\ell)/\ell$ from a horizontal line (see Fig. 10). In particular, the apparent 'oscillation' of this quantity could actually be real. Note that the contribution of the 'noise' term D_η turns out to be of the same order of magnitude as the impact contribution, Eq. (20). This is in fact compatible with the result of Fig. 7.

Note that the purpose of Fig. 9 and 10 is not to fit precisely the data, but only to illustrate an overall agreement between our micro-model and the empirical results. The most striking aspect, however, is that the marginality criterion relating β and γ seems to be quite accurately obeyed.

Coming back to the Fluctuation-Response relation discussed in Section 2.4, we see that our model predicts, for $\ell \gg 1$ where the effect of D_0 can be neglected:

$$\frac{\mathcal{D}(\ell)}{\ell} = Z\langle \ln V\rangle^2 C_0 R_0^2 + D_\eta, \qquad \mathcal{R}(\ell) = Z'\langle \ln V\rangle R_0 C_0, \tag{26}$$

where Z, Z' are numerical constants. Assuming that from one day to the next both the average (log-)traded volume and the impact R_0 of each individual trade

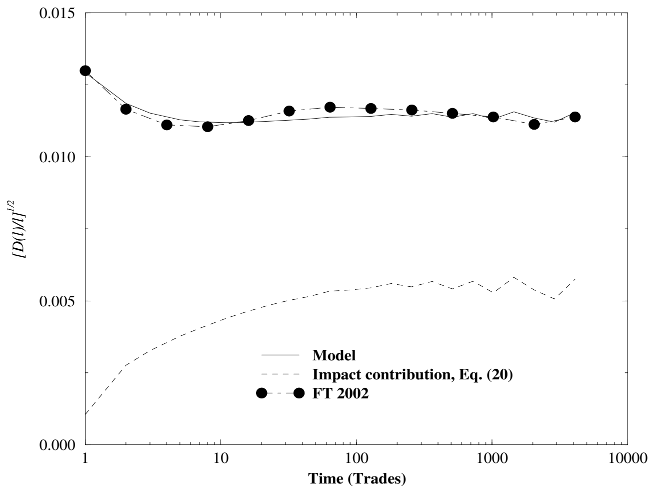

Figure 10: Diffusion constant $\mathcal{D}(\ell)/\ell$, using Eq. (20), with the values of γ, β, C_0, ℓ_0 and R_0 determined from $\mathcal{R}(\ell)$. Two extra parameters were used: $D_\eta = 10^{-4}$ and $D_0 = 6.6\,10^{-5}$ (both in Euro squared).

might change, while C_0 is fixed, immediately leads to the affine relation between $\mathcal{D}$ and $\mathcal{R}^2$ reported in Section 2.4.

The conclusion of this Section is that our 'micro-model' of prices, Eq. (12), can be used as a theoretical canvas to rationalize and interpret the empirical results found in the previous Section. Most surprising is the constraint that these empirical results impose on the shape of the 'bare' response function G_0, which is found to be a slowly decaying power law which must precisely cancel the slowly decaying autocorrelation of the trades. The fact that the bare impact function decays with time (at least on intra-day time scales), in a finely tuned way to compensate the long memory in the trades, is the central result of this paper. This effect is completely lost in zero intelligence models of order flows, where the impact of each trade is permanent ($G_0(\ell) \equiv R_0$), instead of relaxing with time. In fact, both the long time memory of the trades and the slowly relaxing impact function reported here must be the consequence of the strategic behaviour of market participants, that we now discuss in order to get an intuitive understanding of the mechanisms at play.

4 Critical balance of opposite forces: Market orders vs. limit orders

Although trading occurs for a large variety of reasons, it is useful to recognize that traders organize in two broad categories.

- One is that of 'liquidity takers', that trigger trades by putting in market orders. The motivation for this category of traders might be to take advantage of some 'information', and make a profit from correctly anticipating future price changes. Information can in fact be of very different nature: fundamental (firm based), macro-economical, political, statistical (based on regularities of price patterns), etc. Unfortunately, information is often hard to interpret correctly, and it is probable that many of these information driven trades are misguided. For example, systematic hedge funds which take decisions based on statistical pattern recognition have a typical success rate of only 52%. There is no compelling reason to believe that the intuition of traders in markets room fares much better than that. Since market orders allows one to be immediately executed, many impatient investors, who want to liquidate their position, or hedge, etc. might be tempted to place market orders, even at the expense of the bid-ask spread $s(t) = a(t) - b(t)$.

- The other category is that of 'liquidity providers' (or 'market makers', although on electronic markets all participants can be liquidity providers by putting in limit orders), who offer to buy or to sell but avoid taking any bare position on the market. Their profit comes from the bid-ask spread s: the sell price is always slightly larger than the buy price, so that each round turn operation leads to a profit equal to the spread s, at least if the midpoint has not changed in the mean time (see below).

This is where the game becomes interesting. Assume that a liquidity taker wants to buy, so that an increased number of buy orders arrive on the market. The liquidity providers is tempted to increase the offer (or ask) price a because the buyer might be informed and really know that the current price is too low and that it will most probably increase in the near future. Should this happen (and even if, as discussed below) this is in fact quite rare, the liquidity provider, who has to close his position later, might have to buy back at a much higher price. In order not to trigger a sudden increase of a that would make their trade costly, liquidity takers need to put on not too large orders. This is the rationale for dividing one's order in small chunks and disperse these as much as possible over time so as not to appear on the 'radar screens'. Doing so, however, liquidity takers *necessarily create some temporal correlations* in the sign of the trades. Since these traders probably have a somewhat broad spectrum of trading horizons (from a few minutes to several weeks), this can easily explain the slow, power-law decay of the sign correlation function reported above.

Now, if the market orders in fact do *not* contain useful information but are the result of hedging, noise trading, misguided interpretations, errors, etc., then the price should not move up on the long run, and should mean revert to its previous value. Liquidity providers are obviously the active force behind this mean reversion, again because closing their position will be costly if the price has moved up too far from the initial price. A computation of the liquidity provider average gain per share $\mathcal{G}$ can be performed, and is found to be, for trades of volume V [15]:

$$\mathcal{G} = s + \mathcal{R}(0, V) - \mathcal{R}(\infty, V) \approx s + \ln V \left[\mathcal{R}(0) - \mathcal{R}(\infty)\right], \qquad (27)$$

where $\mathcal{R}(0, V)$ is the immediate average impact of a trade, before new limit orders set in. We have in fact checked that $\mathcal{R}(0, V) \approx \mathcal{R}(1, V)$. From the above formula, one sees that it is in the interest of liquidity providers to mean revert the price, such as to make $\mathcal{R}(\infty)$ as small as possible. However, this mean reversion cannot take place too quickly, again because a really informed trader would then be able to buy a large volume at a modest price. Hence, this mean reversion must be slow. From the quantitative analysis of Section 3, we have found that there is hardly any mean reversion at all on short time scales $\ell < \ell_0$, and that this effect can be described as a slow power-law for larger ℓ's. Actually, the action of liquidity providers and liquidity takers must be such that no (or very little) linear correlation is created in the price changes, otherwise one of the two population would statistically make money over the other.

To summarize: liquidity takers must dilute their orders and create long range correlations in the trade signs, whereas liquidity providers must correctly handle the fact that liquidity takers might either possess useful information (a rare situation that can be very costly), but might also be not informed at all and trade randomly. By slowly mean reverting the price, market makers minimize the probability that they either sell too low, or have to buy back too high. The delicate balance between these conflicting tendencies conspire to put the market at the border between persistence (if mean reversion is too weak, i.e. $\beta < \beta_c$) or anti-persistence (if mean reversion is too strong, i.e. $\beta > \beta_c$).

It is actually enlightening to propose a simple model that could explain how market makers enforce this mean reversion. [We have in fact directly checked on the data that the evolution of the midpoint between trades (resulting from the order flow) is anticorrelated with the impact of the trades]. Assume that upon placing limit orders, there is a systematic bias toward some moving average of past prices. If this average is for simplicity taken to be an exponential moving average, the continuous time description of this will read:

$$\begin{aligned} \frac{dp_t}{dt} &= -\Omega(p_t - \overline{p}_t) + \eta_t \\ \frac{d\overline{p}_t}{dt} &= \kappa(p_t - \overline{p}_t), \end{aligned} \qquad (28)$$

where η_t is the random driving force due to trading, Ω the inverse time scale for the strength of the mean reversion, and $1/\kappa$ the 'memory' time over which the average price $\overline{p}_t$ is computed. The first equation means that liquidity providers tend to mean revert the price toward $\overline{p}_t$, while the second describes the update of the exponential moving average $\overline{p}_t$ with time. This set of linear equations can be solved, and leads to a solution of the form $p_t = \int^t dt' G_0(t-t')\eta'_t$, with a bare propagator given by:

$$G_0(t) = (1 - G_\infty)\exp[-(\Omega + \kappa)t] + G_\infty, \tag{29}$$

i.e. an exponential decay toward a finite asymptotic value $G_\infty = \kappa/(\Omega + \kappa)$. (Note that, interestingly, it is the self-referential effect that leads to a non zero asymptotic impact. In the limit where $\kappa \gg \Omega$, the last price is taken as the reference price, and $G_\infty \to 1$). A way to obtain $G_0(t)$ to resemble a power-law is to assume that different market makers use different time horizons to compute a reasonable reference price. This leads to a $G_0(t)$ which writes as the sum of time exponentials with different rates which can easily mimic a pure power-law.

The message of the above model is actually quite interesting from the point of view of informationally Efficient Markets: it suggests that nobody knows what the correct reference price should really be, and that its best proxy is in fact its own past average over some time window (the length of which being itself distributed over several time scales). In the same spirit, the fact that the bare impact function decays as $R_0(\ell_0 + \ell)^{-\beta}$ shows that the information content of each trade is very weak: after 10000 trades (one trading day for FT), and for $\ell_0 = 20$, the impact of a given trade has decreased by a factor $500^{2/5} \approx 12$ as compared to the initial impact. Therefore, the market itself deems that the real information contained in each trade is slim. It would be very interesting to know whether the bare response function levels off to a finite value for large time lags; this will require to go beyond the analysis of the present paper and to deal with overnight effects to enlarge the available range of ℓ values. However, it seems reasonable to expect that $G_0(\ell)$ should indeed reach a finite asymptotic value for values of ℓ corresponding to one or two days of trading.

5 Summary and Conclusion

The aim of this paper was to study in details the statistics of price changes at the trade by trade level, and to analyze the interplay between the impact of each trade on the price and the volatility. Empirical data shows that (a) the price (midpoint) process is close to being purely diffusive, even at the trade by trade scale (b) the temporal structure of the impact function first increases and reaches a maximum after $100 - 1000$ trades, before decreasing back, with a rather limited overall variation (typically a factor 2) and (c) the sign of the trades shows surprisingly long range (power-law) correlations. The paradox is that if the impact of each trade was permanent (as it should be if each trade was informationally efficient) the price process should be strongly super-diffusive and the average response function should increase by a large factor as a function of the time-lag.

As a possible resolution of this paradox, we have proposed a micro-model of prices, Eq. (12) where the price at any instant is the causal result of all past trades, mediated by what we called a bare impact function, or propagator G_0. All the empirical results can be reconciled if one assumes that this bare propagator also decays as a power-law in time, with an exponent which is precisely tuned to a critical value, ensuring simultaneously that prices are nearly diffusive and that the response function is nearly constant. Therefore, the seemingly trivial diffusive behaviour of price changes in fact results from a fined-tuned competition between two opposite effects, one leading to super-diffusion (the autocorrelation of trades) and the other leading to sub-diffusion (the decay of the bare impact function). In financial terms, this competition is between liquidity takers, that create long range correlations by dividing their trading volume in small quantities, and liquidity providers that tend to mean revert the price such as to optimize their gains (see Eq. (27)). As a by product of our micro-model, we justify the fluctuation-response relation suggested by Rosenow [32].

An interesting consequence of our methodology is that one can measure the average information content of each trade, and directly test the premises of the Efficient Market Hypothesis. We have found two quantities that suggest that this content is, perhaps surprisingly, quite small. The decay of the bare impact function over time is one of them; assuming a power-law form for G_0 allows one to quantify this decay, but the conclusion is independent of this particular parameterization: the only way to avoid a large increase of the average response function due to the long range correlations of trades is to have a decaying bare response function. The ratio of the asymptotic value of the bare response function to its initial value is an objective measure of the information content of each trade, and is found to be small. An other, model independent indicator is the asymmetry of the probability distribution of the *signed* price variation, where the sign is that of the trade at the initial time. Information triggered trades should reveal in a detectable positive skew of this distribution, in particular in the tails. Empirical results only show a very small asymmetry, which means that only a small fraction of trades can a posteriori described as truly 'informed', whereas most trades can be classified as noise. This result is most probably one of the mechanism needed to explain the excess volatility puzzle first raised by Schiller [3].

From a more general standpoint, our finding that the diffusive nature of price changes is a result of a critical balance between competing effects is quite interesting and might justify several claims made in the physics literature that the anomalies in price statistics (fat tails in returns described by power laws [19, 20], and long range self similar volatility correlations [23, 24]) are due to the presence of a critical point in the vicinity of which the market operates (see e.g. [34], and in the context of financial markets [35, 36]). If a fine-tuned balance between two competing effects is needed to ensure that on average prices are diffusive, one should expect that fluctuations are crucial, since a local unbalance between the competing forces can lead to an instability. In more financial terms, this is a *liquidity crisis*: a sudden cooperativity of market orders, that lead to an increase of the trade sign correlation function, can out-weight the liquidity providers stabilizing (mean-reverting) role, and lead to crashes. This suggests that one should be able to write a mathematical model, inspired by our results, to describe this 'on-off intermittency' scenario, advocated (although in a different context) in [18, 37, 38].

Acknowledgments: We thank Jelle Boersma, Lisa Borland and Bernd Rosenow for inspiring remarks, and J. Doyne Farmer for communicating his results on the autocorrelation of trades before publication [16] and for insightful discussions. We also thanks S. Picozzi for an interesting discussion and for pointing out the possible relevance of ref. [18] to financial markets. Finally, we take the opportunity of this paper to acknowledge the countless efforts of Gene Stanley to investigate the dynamics of complex systems and to bring together different fields and ideas – as testified by the papers cited in reference, which inspired the present work.

References

[1] E. F. Fama, *Efficient capital markets: A review of theory and empirical work*, Journal of Finance, **25**, 383 (1970).

[2] P. A. Samuelson, *Proof that properly anticipated prices fluctuate randomly*, Industrial Management Review, **6**, 41 (1965).

[3] R. J. Schiller, *Do Stock Prices move too much to be justified by subsequent changes in dividends ?*, American Economic Review, **71**, 421 (1981). R. J. Schiller, *Irrational Exuberance*, Princeton University Press (2000).

[4] W. B. Arthur, *Complexity in Economic and Financial Markets*, Complexity, **1**, 1 (1995).

[5] A. Shleifer, *Inefficient Markets, An Introduction to Behavioral Finance*, Oxford University Press (2000).

[6] A. Orléan, *Le pouvoir de la finance*, Odile Jacob, Paris (1999); *A quoi servent les marchés financiers ?*, in Qu'est-ce que la Culture ?, Odile Jacob, Paris (2001).

[7] P. Bak, M. Paczuski, and M. Shubik, *Price variations in a stock market with many agents*, Physica A **246**, 430 (1997)

[8] S. Maslov, *Simple model of a limit order-driven market*, Physica A **278**, 571 (2000); S. Maslov, M. Millis, *Price fluctuations from the order book perspective – empirical facts and a simple model*, Physica A **299**, 234 (2001)

[9] D. Challet, R. Stinchcombe, *Analyzing and modelling 1+1d markets*, Physica A **300**, 285 (2001)

[10] F. Slanina, *Mean-field approximation for a limit order driven market model*, Phys. Rev. **E 64**, 056136 (2001).

[11] M. G. Daniels, J. D. Farmer, G. Iori, E. Smith, *Quantitative model of price diffusion and market friction based on trading as a mechanistic random process*, Phys. Rev. Lett. **90**, 108102 (2003).

[12] J.P. Bouchaud, M. Mézard, M. Potters, *Statistical properties of stock order books: empirical results and models*, Quantitative Finance **2**, 251 (2002).

[13] E. Smith, J. D. Farmer, L. Gillemot, S. Krishnamurthy, *Statistical theory of the continuous double auction*, e-print cond-mat/0210475, to appear in Quantitative Finance.

[14] D. Challet, R. Stinchcombe, *Non-constant rates and over-diffusive prices in a simple model of limit order markets*, Quantitative Finance **3**, 155 (2003).

[15] M. Wyart, J.P. Bouchaud, in preparation.

[16] J. D. Farmer, talk presented at the ICBI Global Derivatives Trading Summit, Barcelona, May 2003; J. D. Farmer, F. Lillo, in preparation.

[17] C. K. Peng, J. Mietus, J. Hausdorff, S. Havlin, H. E. Stanley, and A. L. Goldberger, *Long-Range Anticorrelations and Non-Gaussian Behavior of the Heartbeat*, Phys. Rev. Lett. 70, 1343-1346 (1993); P. Bernaola-Galvan, P. Ch. Ivanov, L. A. N. Amaral, and H. E. Stanley, *Scale Invariance in the Nonstationarity of Human Heart Rate* Phys. Rev. Lett. **87**, 168105 (2001). The qualitative analogy with financial markets was recently discussed in: Z. Struzik, *Taking the pulse of the economy*, Quantitative Finance, to appear (August 2003).

[18] J. L. Cabrera and J. G. Milton, *On-Off Intermittency in a Human Balancing Task*, Phys. Rev. Lett. 89, 158702 (2002)

[19] T. Lux, *The stable Paretian hypothesis and the frequency of large returns: an examination of major German stocks*, Applied Financial Economics, **6**, 463, (1996).

[20] V. Plerou, P. Gopikrishnan, L.A. Amaral, M. Meyer, H.E. Stanley, *Scaling of the distribution of price fluctuations of individual companies*, Phys. Rev. **E60** 6519 (1999); P. Gopikrishnan, V. Plerou, L. A. Amaral, M. Meyer, H. E. Stanley, *Scaling of the distribution of fluctuations of financial market indices*, Phys. Rev. **E 60** 5305 (1999).

[21] V. Plerou, P. Gopikrishnan, L. A. N. Amaral, X. Gabaix, and H. E. Stanley, *Diffusion and Economic Fluctuations*, Phys. Rev. E **62**, 3023 (2000).

[22] R. Mantegna, H. E. Stanley, *An Introduction to Econophysics*, Cambridge University Press, Cambridge, 1999; J.-P. Bouchaud and M. Potters, *Theory of Financial Risks and Derivative Pricing*, Cambridge University Press, Second Edition, 2003.

[23] for a recent review, see: R. Cont, Quantitative Finance, *Empirical properties of asset returns: stylized facts and statistical issues*, **1**, 223 (2001), and refs. therein.

[24] J.-F. Muzy, J. Delour, E. Bacry, *Modelling fluctuations of financial time series: from cascade process to stochastic volatility model*, Eur. Phys. J. B **17**, 537-548 (2000); for a recent review, see: J.P. Bouchaud, J.F. Muzy, *Financial time series: from Bachelier's random walk to multifractal 'cascades'*, to appear in 'Kolmogorov's Legacy', Editions Belin, Paris 2003.

[25] M. Wyart, J.P. Bouchaud, *Self-referential behaviour, overreaction and conventions in financial markets*, e-print cond-mat/0303584.

[26] F. Black, *Noise*, J. of Finance, **41** 529 (1986).

[27] W. de Bondt, R. Thaler, *Does the market overreact ?*, Journal of Finance, **XL**, 793 (1985).

[28] A. Kempf, O. Korn, *Market Depth and Order Size*, Journal of Financial Markets, **2**, 29 (1999).

[29] V. Plerou, P. Gopikrishnan, X. Gabaix, H. E. Stanley, *Quantifying Stock Price Response to Demand Fluctuations*, Phys. Rev. E **66**, 027104 (2002).

[30] F. Lillo, R. Mantegna, J. D. Farmer, *Single Curve Collapse of the Price Impact Function for the New York Stock Exchange*, e-print cond-mat/0207428; *Master curve for price-impact function*, Nature, **421**, 129 (2003).

[31] M. Potters, J.P. Bouchaud, *More statistical properties of order books and price impact*, Physica A, **324**, 133 (2003).

[32] B. Rosenow, *Fluctuations and Market Friction in Financial Trading*, Int. J. Mod. Phys. C, **13**, 419 (2002). Somewhat related results can also be found in [21]

[33] J.-P. Bouchaud, R. Cont, *A Langevin approach to stock market fluctuations and crashes*, European Journal of Physics, **B 6**, 543 (1998).

[34] P. Bak, *How Nature Works: The Science of Self-Organized Criticality*, Copernicus, Springer, New York, 1996.

[35] D. Challet, A. Chessa, M. Marsili, Y.C. Zhang, *From Minority Games to real markets*, Quantitative Finance, **1**, 168 (2001) and refs. therein; D. Challet, M. Marsili, Y.C. Zhang, *The Minority Game*, book in preparation.

[36] J.P. Bouchaud, *Power-laws in economics and finance: some ideas from physics*, Quantitative Finance, **1**, 105 (2001).

[37] T. Lux, M. Marchesi, *Volatility Clustering in Financial Markets: A Microsimulation of Interacting Agents*, Int. J. Theo. Appl. Fin. **3**, 675 (2000).

[38] A. Krawiecki, J. A. Hoyst, and D. Helbing, *Volatility Clustering and Scaling for Financial Time Series due to Attractor Bubbling*, Phys. Rev. Lett. 89, 158701 (2002).