Topics in market microstructure

Zovko, I.

Publication date
2008

Chapter 2

The power of patience: A behavioral regularity in limit order placement

Published as Ilija I. Zovko, J. Doyne Farmer, *The power of patience: A behavioral regularity in limit order placement*, Quantitative Finance, 2002, No 5, Volume 2, pp: 387-392.

2.1 Introduction

Most modern financial markets are designed as a complex hybrid composed of a continuous double auction and an 'upstairs' trading mechanism serving the purpose of block trades. The double auction is believed to be the primary price discovery mechanism[1]. Limit orders, which specify both a quantity and a limit price (the worst acceptable price), are the liquidity providing mechanism for double auctions and the proper understanding of their submission process is important in the study of price formation.

[1] According to the London Stock Exchange information bulletins ("SETS four years on - October 2001", published by the London Stock Exchange), since the introduction of the SETS in 1997 to October 2001, the average percentage of trades in order book securities that have been executed at the price shown on the order book is 70% - 75%. Therefore SETS seems to serve as the primary price discovery mechanism in London.

We study the *relative limit price* $\delta(t)$, the limit price in relation to the current best price. For buy orders $\delta(t) = b(t) - p(t)$, where p is the limit price, b is the best bid (highest buy limit price), and t is the time when the order is placed. For sell orders $\delta(t) = p(t) - a(t)$, where a is the best ask (lowest sell limit price)[2]. We find a striking regularity in the distribution of relative limit prices and we document clustering of order prices as seen by a slowly decaying autocorrelation function.

Biais, Hillion and Spatt (1995) study the limit order submission process on the Paris Bourse. They note that the number of orders placed up to five quotes away from the market decay monotonically but do not attempt to estimate the distribution or examine orders placed further then five best quotes. Our analysis looks at the price placement of limit orders across a much wider range of prices. Since placing orders out of the market carries execution and adverse selection risk, our work is relevant in understanding the fundamental dilemma of limit order placement: execution certainty vs. transaction costs (see, e.g., Cohen, et al. (1981); Harris (1997); Harris and Hasbrouck (1996); Holden and Chakravarty (1995); Kumar and Seppi (1992); Lo, et al. (2002)).

In addition to the above, our work relates to the literature on clustered volatility. It is well known that both asset prices and quotes display ARCH or GARCH effects (Engle (1982); Bollerslev (1986)), but the origins of these phenomena are not well understood. Explanations range from news clustering (Engle, et al. (1990)), macroeconomic origins (Campbell (1987); Glosten et al. (1993)) to microstructure effects (Lamoureux and Lastrapes (1990); Bollerslev and Domowitz (1991); Kavajecz and Odders-White (2001)). We provide empirical evidence that volatility feedback may in part be caused by limit order placement that in turn depends on past volatility levels.

This paper is organized as follows. Section II introduces the mechanics of limit order trading and describes the London Stock Exchange data we use. Section III presents our results on the distribution and time series properties of relative limit order prices. In section IV we examine the possible relationship of limit order prices and volatility which may lead to volatility clustering. Section V discusses and summarizes the result.

[2]We have made a somewhat arbitrary choice in defining the "best price". An obvious alternative would have been to choose the best ask as the reference price for buy orders, and the best bid as the reference price for sell orders. This would have the advantage that it would have automatically included orders placed inside the interval between the bid and ask (the spread), which are discarded in the present analysis. The choice of reference price does not seem to make a large difference in the tail; for large δ it leads to results that are essentially the same.

2.2 Description of the London Stock Exchange data

The limit order trading mechanism works as follows: As each new limit order arrives, it is matched against the queue of pre-existing limit orders, called the *limit order book*, to determine whether or not it results in any immediate transactions. At any given time there is a best buy price $b(t)$, and a best ask price $a(t)$. A sell order that crosses $b(t)$, or a buy order that crosses $a(t)$, results in at least one transaction. The matching for transactions is performed based on price and order of arrival. Thus matching begins with the order of the opposite sign that has the best price and arrived first, then proceeds to the order (if any) with the same price that arrived second, and so on, repeating for the next best price, etc. The matching process continues until the arriving order has either been entirely transacted, or until there are no orders of the opposite sign with prices that satisfy the arriving order's limit price. Anything that is left over is stored in the limit order book.

On the London Stock Exchange, in addition to limit orders described above, traders can also submit crossing limit orders which result in immediate transactions while limiting market impact. Such crossing limit orders make up about 30% (in the example of Vodafone) of all limit orders and are more like market orders. In this paper we discard them and analyze only limit orders that enter the book. Of the analyzed orders 74% are submitted at the best quotes. Only 1% are submitted inside the spread (with $\delta < 0$), while the remaining 25% are submitted out of the market ($\delta > 0$). We investigate only limit orders with positive relative price $\delta > 0$ and refer to them in text simply as limit orders[3].

The time period of the analysis is from August 1, 1998 to April 31, 2000. This data set contains many errors; we chose the names we analyze here from the several hundred that are traded on the exchange based on the ease of cleaning the data, trying to keep a reasonable balance between high and low volume stocks [4]. This left 50 different names, with a total of roughly seven million limit orders,

[3] Even though limit orders placed in the spread are not numerous, they are very important in price formation. Dataset we use does not include enough events to provide statistically significant results. Our preliminary results indicate that orders placed in the spread behave qualitatively similar to orders placed out of the market, i.e., there are some indications of power law behavior in their limit price density towards the other side of the market.

[4] The ticker symbols for the stocks in our sample are AIR, AL., ANL, AZN, BAA, BARC, BAY, BLT, BOC, BOOT, BPB, BSCT, BSY, BT.A, CCH, CCM, CS., CW., GLXO, HAS, HG., ICI, III, ISYS, LAND, LLOY, LMI, MKS, MNI, NPR, NU., PO., PRU, PSON, RB., RBOS, REED, RIO, RR., RTK, RTO, SB., SBRY, SHEL, SLP, TSCO, UNWS, UU., VOD, and WWH.

of which about two million are submitted out of the market ($\delta > 0$).

2.3 Properties of relative limit order prices

Choosing a relative limit price is a strategic decision that involves a tradeoff between patience and profit (e.g., Holden and Chakravarty (1995); Harris and Hasbrouck (1996); Sirri and Peterson (2002)). Consider, for example, a sell order; the story for buy orders is the same, interchanging "high" and "low". An impatient seller will submit a limit order with a limit price well below $b(t)$, which will typically immediately result in a transaction. A seller of intermediate patience will submit an order with $p(t)$ a little greater than $b(t)$; this will not result in an immediate transaction, but will have high priority as new buy orders arrive. A very patient seller will submit an order with $p(t)$ much greater than $b(t)$. This order is unlikely to be executed soon, but it will trade at a good price if it does. A higher price is clearly desirable, but it comes at the cost of lowering the probability of trading – the higher the price, the lower the probability there will be a trade. The choice of limit price is a complex decision that depends on the goals of each agent. There are many factors that could affect the choice of limit price, such as the time horizon of the trading strategy. *A priori* it is not obvious that the unconditional distribution of limit prices should have any particular simple functional form.

2.3.1 Unconditional distribution

Figure (2.1) shows examples of the cumulative distribution for stocks with the largest and smallest number of limit orders. Each order is given the same weighting, regardless of the number of shares, and the distribution for each stock is normalized so that it sums to one. There is considerable variation in the sample distribution from stock to stock, but these plots nonetheless suggest that power law behavior for large δ is a reasonable hypothesis. This is somewhat clearer for the stocks with high order arrival rates. The low volume stocks show larger fluctuations, presumably because of their smaller sample sizes. Although there is a large number of events in each of these distributions, as we will show later, the samples are highly correlated, so that the effective number of independent samples is not nearly as large as it seems.To reduce the sampling errors we merge the data for all stocks, and estimate the sample distribution for the merged set using the method of ranks, as shown in figure (2.2). We fit the

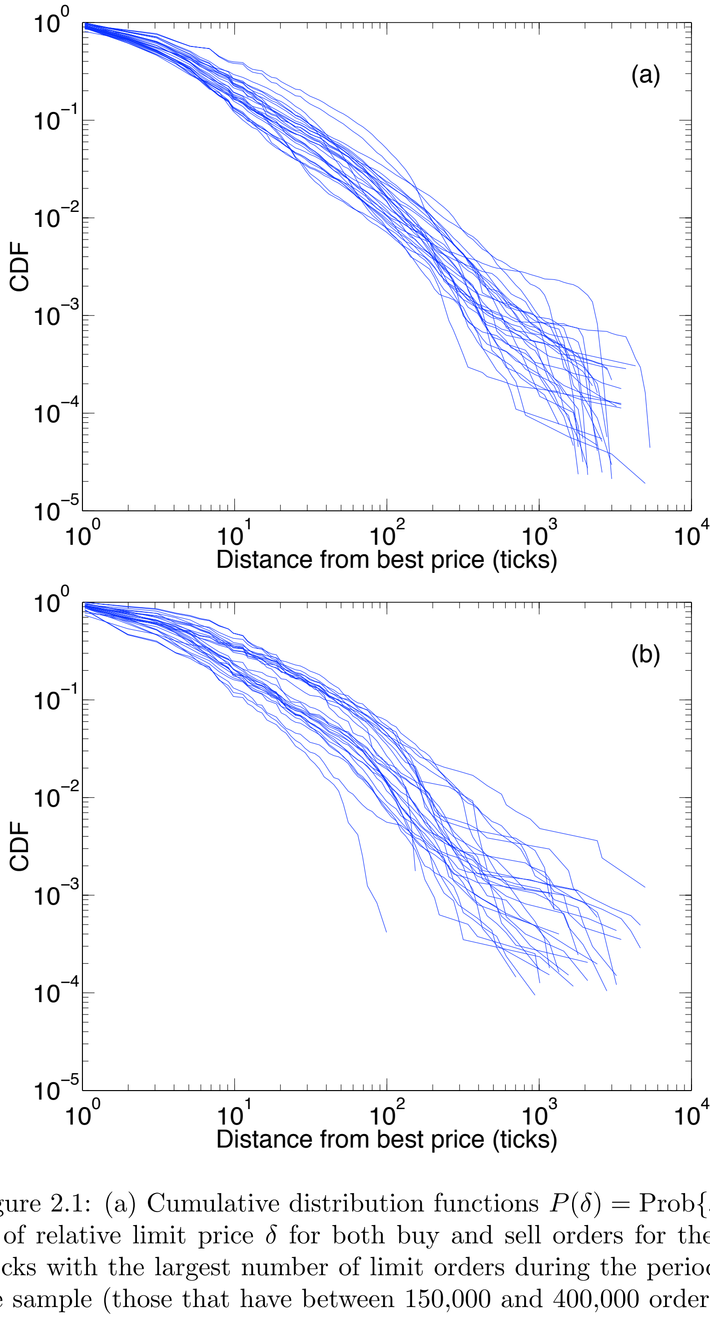

Figure 2.1: (a) Cumulative distribution functions $P(\delta) = \text{Prob}\{x \geq \delta\}$ of relative limit price δ for both buy and sell orders for the 15 stocks with the largest number of limit orders during the period of the sample (those that have between 150,000 and 400,000 orders in the sample.) (b) Same for 15 stocks with the lowest number of limit orders, in the range 2,000 to 100,000. (To avoid overcrowding, we have averaged together nearby bins, which is why the plots appear to violate the normalization condition.)

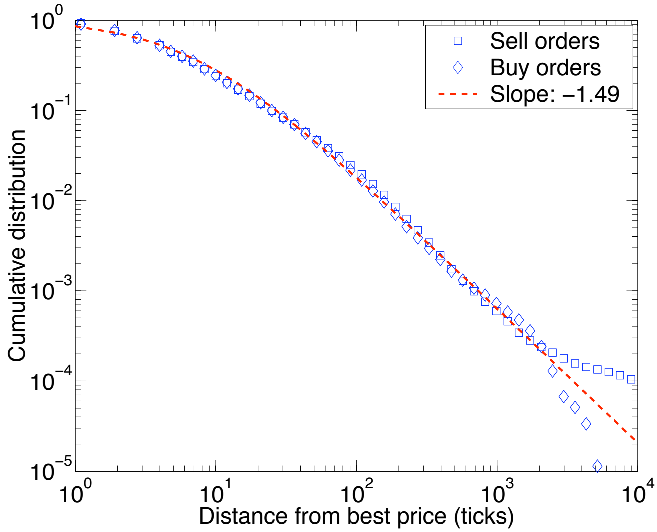

Figure 2.2: An estimate of the cumulative probability distribution based on a merged data set, containing the relative limit order sizes $\delta(t)$ for all 50 stocks across the entire sample. The solid curve is a non-linear least squares fit to the logarithmic form of equation (2.1).

resulting distribution to the functional form[5]

$$P(\delta) = \frac{A}{(x_0 + \delta)^{\beta}}, \tag{2.1}$$

A is set by the normalization, and is a simple function of x_0 and β. Fitting this to the entire sample (both buys and sells) gives $x_0 = 7.01 \pm 0.05$, and $\beta = 1.491 \pm 0.001$. Buys and sells gave similar values for the exponent, i.e. $\beta = 1.49$ in both cases. Since these error bars based on goodness of fit are certainly overly optimistic, we also tested the stability of the results by fitting buys and sells separately on the first and last half of the sample, which gave values in the range $1.47 < \beta < 1.52$. Furthermore, we checked whether there are significant differences in the estimated parameters for stocks with high vs. low order arrival rates. The results ranged from $\beta = 1.5$

[5]The functional form we use to fit the distribution has to satisfy two requirements: it has to be a power law for large δ and finite for $\delta = 0$. A pure power law is either not integrable at 0 or at ∞. If the functional form is to be interpreted as a probability density then it necessarily has to be truncated at one end. In our case the natural truncation point is 0. Clearly there is some arbitrariness in the choice of the exact form, but since we are mainly interested in the behavior for large δ, this functional form seems satisfactory.

for high to $\beta = 1.7$ for low arrival rates, but for the low arrival rate group we do not have high confidence in the estimate.

As one can see from the figure, the fit is reasonably good. The power law is a good approximation across more than two decades, for relative limit prices ranging from about $10 - 2000$ ticks. For British stocks ticks are measured either in pence, half pence, or quarter pence; in the former case, 2000 ticks corresponds to about twenty pounds. Given the low probability of execution for orders with such high relative limit prices this is quite surprising. (For Vodafone, for example, the highest relative limit price that eventually resulted in a transaction was 240 ticks). The value of the exponent $\beta \approx 1.5$ implies that the mean of the distribution exists, but its variance is formally infinite. Note that because normalized power law distributions are scale free, the asymptotic behavior does not depend on units, e.g. ticks vs. pounds. There appears to be a break in the power law at about 2000 ticks, with sell orders deviating above and buy orders deviating below. A break at roughly this point is expected for buy orders due to the fact that $p = 0$ places a lower bound on the limit price. For a stock trading at 10 pounds, for example, with a tick size of a half pence, 2000 ticks is the lowest possible relative limit price for a buy order. The reason for a corresponding break for sell orders is not so obvious, but in view of the extreme low probability of execution, is not surprising. It should also be kept in mind that the number of events in the extreme tail is very low, so this could also be a statistical fluctuation.

2.3.2 Time series properties

The time series of relative limit prices also has interesting temporal structure. This is apparent to the eye, as seen in figure (2.3b), which shows the average relative limit price $\bar{\delta}$ in intervals of approximately 60 events for Barclays Bank. For reference, in figure (2.3a) we show the same series with the order of the events randomized. Comparing the two suggests that the large and small events are more clustered in the real series than in the shuffled series.

This temporal structure appears to be described by a slowly decaying autocorrelation function, as shown in figure (2.4).

One consequence of such a slowly decaying autocorrelation is the slow convergence of sample distributions to their limiting distribution. If we generate artificial IID data with equation (2.1) as its unconditional distribution, the sample distributions converge very quickly with only a few thousand points. In contrast for the real data, even for a stock with 200,000 points the sample distributions display large fluctuations. When we examine subsamples of the real data, the correlations in the deviations across subsamples are ob-

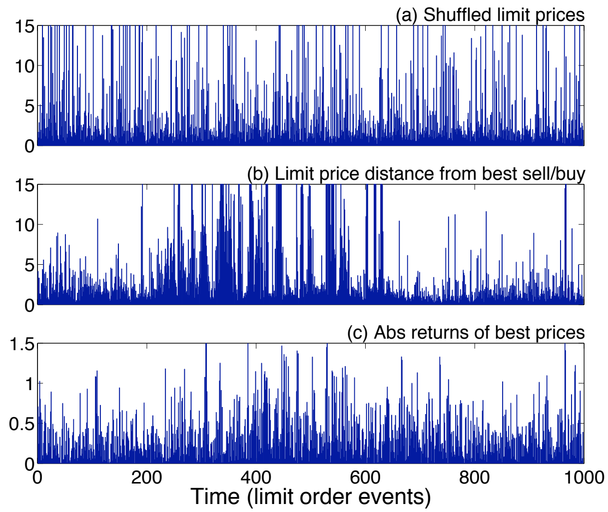

Figure 2.3: (a) Time series of randomly shuffled values of $\delta(t)$ for stock Barclays Bank. (b) True time series $\delta(t)$. (c) The absolute value of the change in the best price between each event in the $\delta(t)$ series.

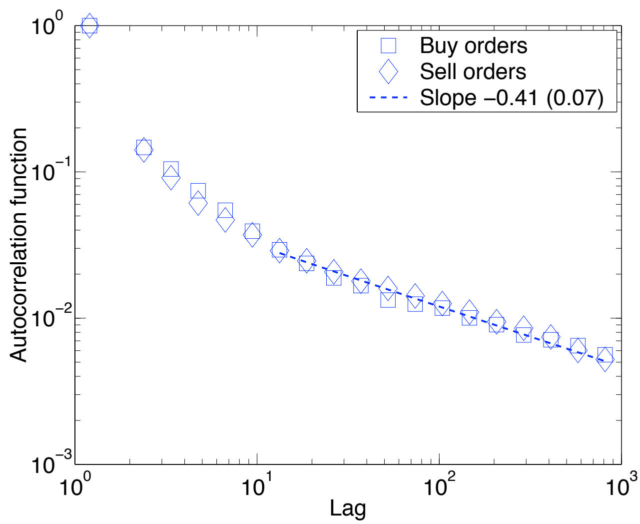

Figure 2.4: The autocorrelation of the time series of relative limit prices δ, averaged across all 50 stocks in the sample, and smoothed across different lags. This is computed in tick time, i.e., the x-axis indicates the number of events, rather than a fixed time.

vious and persist for long periods in time, even when there is no overlap in the subsamples. We believe that the slow convergence of the sample distributions is mainly due to the long range temporal dependence in the data.

2.4 Volatility clustering

To get some insight into the possible cause of the temporal correlations, we compare the time series of relative limit prices to the corresponding price volatility. The price volatility is measured as $v(t) = |\log(b(t)/b(t-1))|$, where $b(t)$ is the best bid for buy orders or the best ask for sell limit orders. We show a typical volatility series in figure (2.3c). One can see by eye that epochs of high limit price tend to coincide with epochs of high volatility.

To help understand the possible relation between volatility and relative limit price we calculate their cross-autocorrelation. This is defined as

$$XCF(\tau) = \frac{\langle v(t-\tau)\delta(t)\rangle - \langle v(t)\rangle\langle\delta(t)\rangle}{\sigma_v\sigma_\delta}, \tag{2.2}$$

where $\langle\cdot\rangle$ denotes a sample average, and σ denotes the standard deviation. We first create a series of the average relative limit price and average volatility over 10 minute intervals. We then compute the cross-autocorrelation function and average over all stocks. The result is shown in figure (2.5).

We test the statistical significance of this result by testing against the null hypothesis that the volatility and relative limit price are uncorrelated. To do this we have to cope with the problem that the individual series are highly autocorrelated, as demonstrated in figure (2.4), and the 50 series for each stock also tend to be correlated to each other. To solve these problems, we construct samples of the null hypothesis using a technique introduced in Theiler, et al (1992). We compute the discrete Fourier transform of the relative limit price time series. We then randomly permute the phases of the series, and perform the inverse Fourier transform. This creates a realization of the null hypothesis, drawn from a distribution with the same unconditional distribution and the same autocorrelation function. Because we use the same random permutation of phases for each of the 50 series, we also preserve their correlation to each other. We then compute the cross autocorrelation function between each of the 50 surrogate limit price series and its corresponding true volatility series, and then average the results. We then repeat this experiment 300 times, which gives us a distribution of realizations of averaged sample cross-autocorrelation functions under the null

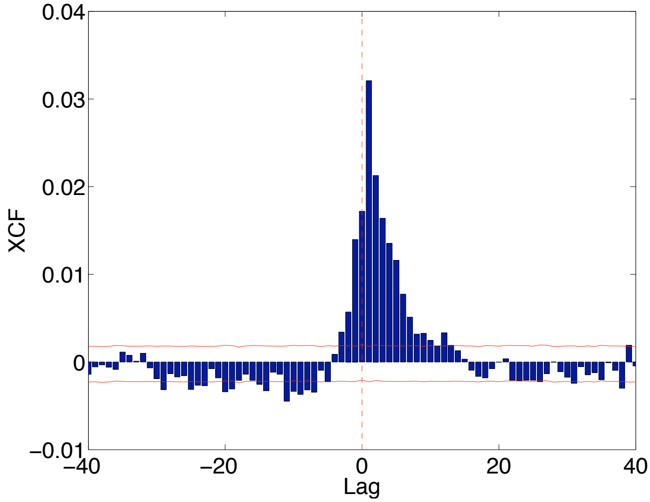

Figure 2.5: The cross autocorrelation of the time series of relative limit prices $\delta(t)$ and volatilities $v(t-\tau)$, averaged across all 50 stocks in the sample.

hypothesis. This procedure is more appropriate in this case than the standard moving block bootstrap, which requires choice of a timescale and will not work for a series such as this that does not have a characteristic timescale. The 2.5% and 97.5% quantile error bars at each lag are denoted by the two solid lines near zero in figure (2.5).

From this figure it is clear that there is indeed a strong contemporaneous correlation between volatility and relative limit price, and that the result is highly significant. Furthermore, there is some asymmetry in the cross-autocorrelation function; the peak occurs at a lag of one rather than zero, and there is more mass on the right than on the left. This suggests that there is some tendency for volatility to lead the relative limit price. This implies one of three things: (1) Volatility and limit price have a common cause, but this cause is for some reason felt later for the relative limit price; (2) the agents placing orders key off of volatility and correctly anticipate it; or, more plausibly, (3) volatility at least partially causes the relative limit price. Angel (1994) has suggested that volatility might affect limit order placement in this way.

Note that this suggests an interesting feedback loop: Holding other aspects of the order placement process constant, an increase in the average relative limit price will lower the depth in the limit order book at any particular price level, and therefore increase volatility. Since such a feedback loop is unstable, there are presumably nonlinear feedbacks of the opposite sign that eventually damp it. Nonetheless, such a feedback loop may potentially contribute to creating clustered volatility.

2.5 Conclusion

One of the most surprising aspects of the power law behavior of relative limit price is that traders place their orders so far away from the current price. As is evident in figure (2.2), orders occur with relative limit prices as large as 10,000 ticks (or 25 pounds for a stock with ticks in quarter pence). While we have taken some precautions to screen for errors, such as plotting the data and looking for unreasonable events, despite our best efforts, it is likely that there are still data errors remaining in this series. There appears to be a break in the merged unconditional distribution at about 2000 ticks; if this is statistically significant, it suggests that the very largest events may follow a different distribution than the rest of the sample, and might be dominated by data errors. Nonetheless, since we know that most of the smaller events are real, and since we see no break in the behavior until roughly $\delta \approx 2000$, errors are highly unlikely to be the cause of the power law behavior seen for $\delta < 2000$.

The conundrum of very large limit orders is compounded by consideration of the average waiting time for execution as a function of relative limit price. We intend to investigate the dependence of the waiting time on the limit price in the future, but since this requires tracking each limit order, the data analysis is more difficult. We have checked this for one stock, Vodafone, in which the largest relative limit price that resulted in an eventual trade was $\delta = 240$ ticks. Assuming other stocks behave similarly, this suggests that either traders are strongly over-optimistic about the probability of execution, or that the orders with large relative limit prices are placed for other reasons.

Since obtaining our results we have seen a recent preprint by Bouchaud et al. (2002) analyzing three stocks on the Paris Bourse over a period of a month. They also obtain a power law for $P(\delta)$, but they observe an exponent $\beta \approx 0.6$, in contrast to our value $\beta \approx 1.5$. We do not understand why there should be such a discrepancy in results. While they analyze only three stock-months of data, whereas we have analyzed roughly 1050 stock-months, their order arrival rates are roughly 20 times higher than ours, and their sample distributions appear to follow the power law scaling fairly well.

One possible explanation is the long-range correlation. Assuming the Paris data show the same behavior we have observed, the decay in the autocorrelation is so slow that there may not be good convergence in a month, even with a large number of samples. The sample exponent $\hat{\beta}$ based on one month samples may vary with time, even if the sample distributions appear to be well-converged. It is of course also possible that the French behave differently than the British, and that for some reason the French prefer to place orders much further from the midpoint.

Our original motivation for this work was to model price formation in the limit order book, as part of the research program for understanding the volatility and liquidity of markets outlined in Daniels, et al. (2001). $P(\delta)$ is important for price formation, since where limit orders are placed affects the depth of the limit order book and hence the diffusion rate of prices. The power law behavior observed here has important consequences for volatility and liquidity that will be described in a future paper.

Our results here are interesting for their own sake in terms of human psychology. They show how a striking regularity can emerge when human beings are confronted with a complicated decision problem. Why should the distribution of relative limit prices be a power law, and why should it decay with this particular exponent? Our results suggest that the volatility leads the relative limit price, indicating that traders probably use volatility as a signal when placing orders. This supports the obvious hypothesis that traders are reasonably aware of the volatility distribution when placing orders, an effect that may contribute to the phenomenon of clustered volatility. Plerou et al. (1999) have observed a power law for the unconditional distribution of price fluctuations. It seems that the power law for price fluctuations should be related to that of relative limit prices, but the precise nature and the cause of this relationship is not clear. The exponent for price fluctuations of individual companies reported by Plerou et al. is roughly 3, but the exponent we have measured here is roughly 1.5. Why these particular exponents? Makoto Nirei has suggested that if traders have power law utility functions, under the assumption that they optimize this utility, it is possible to derive an expression for β in terms of the exponent of price fluctuations and the coefficient of risk aversion. However, this explanation is not fully satisfying, and more work is needed. At this point the underlying cause of the power law behavior of relative limit prices remains a mystery.