The position profiles of order cancellations in an emerging stock market

Gao-Feng Gu\textsuperscript{1,2}, Xiong Xiong\textsuperscript{3}, Fei Ren\textsuperscript{1,2,4}, Wei-Xing Zhou\textsuperscript{1,2,4} and Wei Zhang\textsuperscript{3}

\textsuperscript{1} School of Business, East China University of Science and Technology, Shanghai 200237, People’s Republic of China
\textsuperscript{2} Research Center for Econophysics, East China University of Science and Technology, Shanghai 200237, People’s Republic of China
\textsuperscript{3} College of Management and Economics, Tianjin University, Tianjin 300072, People’s Republic of China
\textsuperscript{4} School of Science, East China University of Science and Technology, Shanghai 200237, People’s Republic of China

E-mail: gfgu@ecust.edu.cn, xxpeter@tju.edu.cn, fren@ecust.edu.cn, wxzhou@ecust.edu.cn and weiz@tju.edu.cn

Received 3 November 2012
Accepted 22 March 2013
Published 19 April 2013

Abstract. Order submission and cancellation are two constituent actions of stock trading behaviors in order-driven markets. Order submission dynamics has been extensively studied for different markets, while order cancellation dynamics is less understood. There are two positions associated with a cancellation, that is, the price level in the limit-order book (LOB) and the position in the queue at each price level. We study the profiles of these two order cancellation positions through rebuilding the limit-order book using the order flow data of 23 liquid stocks traded on the Shenzhen Stock Exchange in the year 2003. We find that the profiles of relative price levels where cancellations occur obey a log-normal distribution. After normalizing the relative price level by removing the factor of order numbers stored at the price level, we find that the profiles exhibit a power-law scaling behavior on the right tails for both buy and sell orders. When focusing on the order cancellation positions in the queue at each price level, we find that the profiles increase rapidly in the front of the queue, and then fluctuate around a constant value till the end of the queue. These profiles are similar for different stocks. In addition, the profiles of cancellation positions can...
be fitted by an exponent function for both buy and sell orders. These two kinds of cancellation profiles seem universal for different stocks investigated and exhibit minor asymmetry between buy and sell orders. Our empirical findings shed new light on the order cancellation dynamics and pose constraints on the construction of order-driven stock market models.

**Keywords:** critical phenomena of socio-economic systems, scaling in socio-economic systems

---

### Contents

1. Introduction 2
2. Data sets 4
3. The profile of cancellation positions in the limit-order book 7
4. The profile of cancellation positions at the price level 12
5. Conclusion 16
   - Acknowledgments 17
   - References 17

---

### 1. Introduction

In order-driven markets, order cancellation plays an important role in price formation and makes a significant contribution to the order book activity. When limit orders at the best price are cancelled completely, mid-price (the mean of best bid and best ask) and spread will change as well. If cancellation occurs inside the limit-order book (LOB), it also affects the shape of LOB and has potential effects on price formation.

The motivation of order cancellation is to avoid risk, including non-execution (NE) risk and free option (FO) risk. The former risk arises when the current price moves away from the submitting price, so limit orders may not be executed immediately. NE risk may cause traders to suffer opportunity cost [1]. In order to avoid NE risk, traders may submit more aggressive limit orders or market orders to increase the execution probability. FO risk arises when good or bad news arrives, which drives the intrinsic value of the asset to be underestimated or overestimated based on the current market price. Traders cancel their stale orders in time to prevent them from being traded at unfavorable prices. Thus, if traders are concerned with the NE risk, they will revise their orders to increase the price priority, while if traders focus on the FO risk, they will revise orders to decrease the price priority. In the Australian security market, NE risk is the major reason for order cancellation, especially for large orders [2].

For the reason that order cancellation concerns the dynamics of LOB and there are not enough cancellation flow data in the past, only a few papers refer to the study of order

doi:10.1088/1742-5468/2013/04/P04027
cancellation. However, in recent years, with the development of information technology and the rapid development of electronic trading venues, traders can easily conduct an order cancellation at their private computers, which has led to a striking increase of order cancellation in financial markets. For example, more than 20% of submitted orders are cancellation orders, and 40% of limit orders are cancelled on the New York Stock Exchange [3]. The cancellation percentage is even higher in futures markets [4], where the cancellation rate is as high as 68.3%.

Many scholars have paid attention to the study of order cancellation in recent years and presented useful results. Yeo [3] studied a database of 148 stocks traded on the New York Stock Exchange in the year 2001 and found that about 95% of order cancellations occur within 10 min after submission. He also reported that traders tend to resubmit limit orders with more aggressive prices after a cancellation to avoid NE risk. Fong and Liu [2] studied the order flow data of 40 stocks listed on the Australian Securities Exchange (ASX) in August 2000. They found that half of the cancellations occur within the first two price levels in the LOB, and indicated that cancellation probability increases with order size, but decreases with order aggressiveness. Liu [5] presented a simple model for order revision and cancellation. The paper shows that cancellation is positively related to order submission risks, and that large capitalization stocks tend to have more cancellations. Hasbrouck and Saar [6] focused on cancellations on the Island electronic communications network (ECN). They observed that a large number of limit orders are cancelled very shortly after their submissions and two sharp jumps appear in the probability density function (PDF) of order cancellation. Biais et al [7] analyzed a database of 40 stocks in the year 1991 on the Paris Bourse, and found that cancellations at the bid or ask side are relatively frequent after large sales or purchases. The paper shows that cancellations follow each other quickly on the same side of the LOB as well. Ellul et al [8, 9] studied a data sample of 148 stocks trading on the NYSE, and found that when a buy or sell order is cancelled, the most likely subsequent event is the arrival of the same kind of limit orders. Ni et al [10] analyzed the statistical properties of inter-cancellation duration (defined as the waiting time between consecutive cancellations) of 22 stocks on the Chinese stock market. They found that the probability density function of durations can be fitted by a Weibull distribution, and that the duration time series exhibits long-term memory and a multifractal nature with the detrended fluctuation analysis (DFA) and the multifractal DFA methods. They reached the conclusion that the order cancellation is a non-Poisson process.

The relationship between cancellation and bid–ask spread has been wildly studied, although the results are controversial. Some scholars concluded that cancellation is positively related to spread [2, 3], others however showed that cancellation is more likely to happen when spread is tight [5, 8, 9]. In addition, order cancellation is more frequent at the opening and closing time on a trading day, displaying a U-shape intra-day pattern [2, 4, 5, 7, 9, 8].

In this paper, we will study the empirical position profiles of order cancellations by rebuilding the LOBs using 23 liquid stocks traded on the Shenzhen Stock Exchange (SZSE) in the Chinese stock market. The rest of paper is organized as follows. In section 2, we will describe briefly the database we adopt. Section 3 presents the cancellation position profile of price levels in the LOB. We further discuss the cancellation position in a queue at a certain price level and study its profile in section 4. Section 5 summarizes the results.
2. Data sets

There are two kinds of auctions held on the Shenzhen Stock Exchange, called opening call auctions and continuous auctions. Opening call auctions are held to generate the opening price at the beginning of a trading day. It corresponds to the process of one-time centralized matching of buy and sell orders accepted during the period from 9:15 am to 9:25 am. Continuous auctions operate from 9:30 am to 11:30 am and 13:00 pm to 15:00 pm. It refers to the process of continuous matching of buy and sell orders on a one-by-one basis. The interval between opening call auctions and continuous auctions (9:25 am–9:30 am) is the cool period, when the exchange is open to orders routing from members, but does not process orders or process cancellations. Closing call auctions (14:57 pm–15:00 pm) were not held until July 1 2006 to generate the closing price.

In an order-driven market, limit-order book is a queue of limit orders waiting to be executed and it is the basis of the continuous double auction mechanism. Limit orders are arranged in the LOB according to the price–time priority, which consists of the price priority principle and the time priority principle. The price priority principle means priority is given to a higher buy order over a lower buy order and a lower sell order is prioritized over a higher sell order. The time priority principle concerns orders having the same submitting price and indicates that priority is given to the order received earliest by the exchange trading system.

Price levels in the LOB are discrete. The difference between two adjacent price levels is the tick size or its multiple. In the Chinese stock market, the tick size is 0.01 CNY for all the A shares of stocks. Figure 1 presents the structure of the LOB. In the figure, we denote $x$ as the price level in the LOB. In the buy LOB, orders with higher limit prices have execution priority and are stored in the front of the LOB, so $x$ decreases with buy price. In contrast, in the sell LOB, priority is given to the orders with lower limit prices, and $x$ increases with the sell price. The highest buy price allocated at the first price level in the buy LOB ($x = 1$) is called the best bid, and the lowest sell price allocated at the first price level in the sell LOB ($x = 1$) is called the best ask. Spread is defined as the difference between the best bid and best ask. When limit orders are submitted with the

---

Figure 1. The structure of the limit-order book.
The position profiles of order cancellations

Table 1. The number and percentage of cancellations of buy and sell orders in the continuous auction for 23 stocks. The columns show the stock code, the number of cancellations $C$ and the ratio $r$ of the number of cancelled buy (or sell) orders to the number of all buy (or sell) orders, respectively.

| Stock | Buy orders | Sell orders |
|-------|------------|-------------|
|       | $C$ | $r$ (%) | $C$ | $r$ (%) |
| 000001 | 317015 | 19.2 | 274929 | 18.3 |
| 000002 | 34577  | 4.2  | 43801  | 4.9  |
| 000009 | 183804 | 21.9 | 187421 | 20.7 |
| 000012 | 114662 | 24.7 | 106361 | 24.3 |
| 000016 | 60219  | 21.1 | 59189  | 19.2 |
| 000021 | 157174 | 23.8 | 152853 | 22.0 |
| 000024 | 42593  | 22.3 | 45594  | 19.6 |
| 000027 | 33058  | 7.1  | 38647  | 5.8  |
| 000063 | 25681  | 7.2  | 34815  | 7.5  |
| 000066 | 110289 | 24.3 | 106695 | 21.8 |
| 000088 | 6861   | 5.6  | 6917   | 4.5  |
| 000089 | 20909  | 7.0  | 22984  | 7.6  |
| 000406 | 94100  | 20.3 | 91776  | 20.1 |
| 000429 | 36999  | 20.1 | 36174  | 17.4 |
| 000488 | 32585  | 18.2 | 33854  | 18.4 |
| 000539 | 26950  | 15.0 | 27087  | 15.9 |
| 000541 | 19715  | 18.8 | 19847  | 16.6 |
| 000550 | 122865 | 23.6 | 129606 | 22.9 |
| 000581 | 27236  | 18.6 | 29490  | 16.2 |
| 000625 | 123361 | 22.2 | 131719 | 22.6 |
| 000709 | 65704  | 18.7 | 64017  | 17.1 |
| 000720 | 16558  | 11.4 | 14209  | 10.7 |
| 000778 | 43576  | 18.4 | 47360  | 16.4 |

same price, they are also stored in a queue at a certain price level according to the arriving time. Denote $y$ as the sequence of orders’ arrival times, thus a smaller value of $y$ indicates an earlier arrival time among the orders with the same submission price.

Our study is based on the order flow data of 23 liquid stocks extracted from the A-share market on the SZSE in 2003. The database contains details of order placement and order cancellation, with the time stamp accurate to 0.01 s. Table 1 depicts the cancellation number $C$ for both buy and sell orders of 23 stocks. We calculate the ratio $r$ of the number of cancelled buy (or sell) orders to the number of all buy (or sell) orders, including effective limit orders placed in the LOB and marketable orders that are fully or partly filled. We find that the ratio $r$ fluctuates within a wide range, with $[4.2\%, 24.7\%]$ for buy orders and with $[4.5\%, 24.3\%]$ for sell orders. There is an interesting feature that the ratio of buy orders is close to sell orders for each stock, which means that a large proportion of buy orders cancelled corresponds to a large number of cancellations taking place for sell orders, and vice versa. For all the stocks, we have mean values $\bar{r} = 16.1\%$ for buy orders and $\bar{r} = 17.1\%$ for sell orders.

doi:10.1088/1742-5468/2013/04/P04027
The cancellation behavior is related to the aggressiveness of submitted orders. There are five types of orders according to their submission aggressiveness, that is, fully filled orders, partially filled orders, orders placed inside the bid–ask spread, orders placed at the same best price, and orders placed inside the LOB. For buy or sell orders of each type, except for fully filled orders that do not have cancellations, we calculate the ratio \( r_1, r_2, r_3, \) and \( r_4 \) of the number cancellations to the number of all orders of the same type and order direction. The result is presented in table 2.

It is shown that the cancellation ratio decreases with the submission aggressiveness. For the 23 stocks analyzed, we obtain the mean values \( r_1 = 0.5\% \), \( r_2 = 11.1\% \), \( r_3 = 18.2\% \) and \( r_4 = 27.7\% \) for buy orders, and the mean values \( r_1 = 0.6\% , r_2 = 10.3\% , r_3 = 19.2\% \) for sell orders.
and \( \tau_4 = 21.3\% \) for sell orders. We find that a very small proportion of partially filled orders have been cancelled and more aggressive orders are less likely to be cancelled. These findings indicate that traders are more likely to cancel orders that have a lower chance to be executed to reduce the NE risk.

3. The profile of cancellation positions in the limit-order book

In this section, we study the profile of cancellation positions (probability distribution of cancellation position) in the LOB. Denote \( x(t) \) as the price level where a cancellation occurred in the LOB at time \( t \). Here we use the event time \( t \), i.e. when a cancellation happens, it increases by one (\( t \to t + 1 \)). Suppose an order marked with gray color in figure 1 is cancelled at time \( t \). The cancelled order is stored at the second price level in the buy LOB, i.e. \( x(t) = 2 \). Generally, \( x(t) = n \) indicates that a cancellation occurs at the \( n \)th price level in the buy or sell LOB at time \( t \).

The length of LOB, \( L_b(t) \) or \( L_s(t) \), defined as the number of price levels existing in the buy or sell LOB, plays an important role in the order cancellation process. For example, it is a completely different situation that a trader cancels an order at the second price level \( x(t) = 2 \) when \( L(t) = 5 \) from the situation that they cancel an order at the second price level when \( L(t) = 50 \). Since it is trivial to study the position profile of the absolute values of \( x(t) \), we will focus on the relative price levels \( X(t) \) instead, which is defined as follows

\[
X(t) = \frac{x(t)}{L_{b,s}(t)},
\]

where \( X(t) \) varies in the range (0, 1]. According to the definition, a small value of \( X \) means that the cancelled order located in the front of the LOB in reference to the best bid or ask price, while a large value of \( X \) suggests that the cancellation happens far from the best price.

We adopt the probability density function (PDF) to describe the position profile of order cancellations, and analyze the PDF \( f(X) \) of relative price levels of cancellations for both buy and sell orders. In figure 2, we illustrate the probability density functions \( f(X) \) for four stocks randomly chosen from the 23 stocks analyzed.

As we know, the order stored in the LOB is executed only when all the orders ahead have been removed, so the orders at the end of the LOB are less likely to be executed and its cancellation probability increases as well. We expect that the probability density function \( f(X) \) should increase monotonically with respect to the price level \( X \). However, as presented in figure 2, the PDFs \( f(X) \) first increase sharply, then drop rapidly for both buy and sell orders, which deviates from our expectation. The special shape of the PDF may be caused by the reason that in the Chinese stock market many impatient traders place aggressive orders close to the same best price to increase the transaction probability, which is confirmed by the fact that the best price orderbook shape has a maximum away from the same best price \( X = 1 \) but extremely close to the same best price [11]. Once the impatient traders cannot make an immediate transaction, they will cancel the placed orders and place a new order to catch the price in order to avoid NE risk, which results in a large cancellation probability in the front of the LOB. In addition, the PDFs \( f(X) \) are not identical between buy orders and sell orders. For sell orders, they have higher
The position profiles of order cancellations

Figure 2.
Probability density functions $f(X)$ of relative price levels of cancellations for both buy and sell orders of four randomly chosen stocks: (a) 000 012, (b) 000 021, (c) 000 066 and (d) 000 0625. The solid lines are fits to the log-normal distribution.

cancellation probability in the front and at the end of the sell LOB, however buy orders have higher cancellation probability in the middle of the buy LOB. The sharply increasing value of $f(X)$ located at $X = 1$ may be caused by the 10% price limit trading rules in the Chinese stock market.

In figure 2, we observe that the PDF $f(X)$ of relative price levels follows a log-normal distribution. Since the relative price $X$ varies in the range $(0, 1]$, we normalize the log-normal distribution to fit the PDF, that is,

$$f(X) = \frac{1}{z} \frac{1}{\sqrt{2\pi\sigma X}} \exp \left[ -\frac{(\ln X - \mu)^2}{2\sigma^2} \right],$$

(2)

where $\mu$ is the location parameter and $\sigma$ is the scaling parameter, and $z$ is a normalized constant ($z = \int_0^1 (1/\sqrt{2\pi\sigma x}) \exp \left[-((\ln x - \mu)^2/2\sigma^2)\right] dx$). We analyze the remaining stocks as well, and find that their PDFs are all log-normally distributed. Using the least squares fitting method, we calculate the parameters $\mu$ and $\sigma$ for both buy and sell orders of 23 stocks, which are listed in table 3. In order to compare the performance of the log-normal distribution fit, we introduce a new statistical variable $rms$ (defined as the root mean square of the difference between the best fit and the empirical data), which is also depicted.
in table 3. We also apply a gamma distribution to fit the data and find that the log-normal distribution gives better results.

According to table 3, the value of $\mu$ obtained from cancelled sell orders is less than cancelled buy orders for each stock, while the value of $\sigma$ of cancelled sell orders is greater than cancelled buy orders. We also calculate the mean values of $\mu$ and $\sigma$ and find that $\overline{\mu} = -2.01 \pm 0.18$ and $\overline{\sigma} = 1.13 \pm 0.09$ for buy orders and $\overline{\mu} = -2.26 \pm 0.15$ and $\overline{\sigma} = 1.34 \pm 0.11$ for sell orders. It means that there exists a higher peak in the PDF for cancelled sell orders and the peak for cancelled sell orders is closer to the same best price, which implies that sellers are more eager to make an immediate transaction. This market behavior is consistent with the situation that it was a bear stock market during the year 2003 and market participants were more willing to sell their orders.

Since the majority of PDFs $f(X)$ of the 23 stocks are log-normally distributed, as shown in table 3, we aggregate all the data together and treat them as an ensemble to obtain a better statistic. As expected, the PDFs of ensemble relative price levels obey the log-normal distribution for both buy and sell orders. Using the same method, we obtain

Table 3. Characteristic parameters of PDFs $f(X)$ for both buy and sell orders of 23 stocks. $\mu$ and $\sigma$ are the parameters of the log-normal distribution obtained by the least squares fitting method. $\text{rms}$ is the root mean square of the difference between the best fit and the empirical data. $\text{Prob}$ is the p-value calculated from the Monte Carlo method with 1000 repeats.

| Stock | Buy orders | Sell orders |
|-------|------------|-------------|
|       | $\mu$    | $\sigma$   | $\text{rms}$ | $\text{Prob}$ | $\mu$    | $\sigma$   | $\text{rms}$ | $\text{Prob}$ |
| 00001 | -2.36    | 1.13       | 0.13         | 0.00         | -2.49    | 1.52       | 0.09         | 0.00         |
| 00002 | -2.17    | 1.10       | 0.33         | 0.89         | -2.37    | 1.22       | 0.11         | 0.00         |
| 00009 | -2.26    | 1.01       | 0.21         | 0.28         | -2.27    | 1.26       | 0.16         | 0.15         |
| 00012 | -2.04    | 1.10       | 0.07         | 0.00         | -2.28    | 1.37       | 0.10         | 0.00         |
| 00016 | -2.00    | 1.07       | 0.16         | 0.01         | -2.26    | 1.27       | 0.11         | 0.00         |
| 00021 | -2.22    | 1.13       | 0.09         | 0.00         | -2.29    | 1.45       | 0.11         | 0.00         |
| 00024 | -1.93    | 1.13       | 0.10         | 0.00         | -2.25    | 1.30       | 0.10         | 0.00         |
| 00027 | -2.14    | 1.04       | 0.10         | 0.00         | -2.26    | 1.23       | 0.13         | 0.00         |
| 00063 | -2.02    | 1.22       | 0.08         | 0.00         | -2.26    | 1.50       | 0.13         | 0.00         |
| 00066 | -2.11    | 1.10       | 0.07         | 0.00         | -2.29    | 1.35       | 0.11         | 0.00         |
| 00088 | -1.58    | 1.21       | 0.25         | 0.68         | -2.06    | 1.43       | 0.08         | 0.00         |
| 00089 | -2.01    | 1.14       | 0.17         | 0.06         | -2.32    | 1.23       | 0.12         | 0.00         |
| 00406 | -2.25    | 1.03       | 0.09         | 0.00         | -2.40    | 1.26       | 0.15         | 0.00         |
| 00429 | -1.99    | 0.98       | 0.18         | 0.02         | -2.33    | 1.18       | 0.09         | 0.00         |
| 00488 | -1.88    | 1.26       | 0.11         | 0.00         | -2.12    | 1.34       | 0.11         | 0.00         |
| 00539 | -1.80    | 1.14       | 0.21         | 0.37         | -2.09    | 1.37       | 0.09         | 0.00         |
| 00541 | -1.80    | 1.23       | 0.18         | 0.20         | -2.40    | 1.39       | 0.15         | 0.04         |
| 00550 | -2.01    | 1.13       | 0.08         | 0.00         | -2.21    | 1.34       | 0.12         | 0.00         |
| 00581 | -1.75    | 1.13       | 0.17         | 0.08         | -2.24    | 1.35       | 0.10         | 0.00         |
| 00625 | -2.07    | 1.18       | 0.10         | 0.00         | -2.18    | 1.38       | 0.13         | 0.04         |
| 00709 | -2.13    | 1.00       | 0.12         | 0.00         | -2.53    | 1.11       | 0.16         | 0.00         |
| 00720 | -1.77    | 1.31       | 0.16         | 0.09         | -1.79    | 1.56       | 0.16         | 0.11         |
| 00778 | -2.02    | 1.11       | 0.12         | 0.00         | -2.33    | 1.38       | 0.12         | 0.01         |
The position profiles of order cancellations

the parameters $\mu = -2.14$ and $\sigma = 1.11$ with $rms = 0.07$ for ensemble buy orders, and $\mu = -2.29$ and $\sigma = 1.35$ with $rms = 0.11$ for ensemble sell orders, which are all close to their mean values mentioned above.

Except for the factor of price levels in the LOB, the number of orders stored in the LOB also affects traders’ cancellation strategy, especially for the number of orders stored at the price level where the cancellation takes place. Denote $n(x, t)$ as the number of orders stored at price level $x$ at time $t$. We conjecture that the cancellation probability of an order at price level $x(t)$ is positively correlated with $n(x, t)$. Our purpose is to analyze the relation between the cancellation probability and the cancellation position in the LOB, and it is necessary to remove the number effect at each price level. We define the normalized relative price levels $\tilde{X}(t)$ by

$$\tilde{X}(t) = \frac{X(t)}{\sum_{x=1}^{L_{b,s}(t)} n(x, t)},$$

where $L_{b,s}(t)$ is the length of the buy or sell LOB and $\sum_{x=1}^{L_{b,s}(t)} n(x, t)$ is the total number of orders stored in the buy or sell LOB at time $t$. This normalization aims at flattening the order book shape, which usually has a hump for stocks in the Chinese market [11] and in other markets as well [12]–[17]. In this way, the numbers of orders at different price levels can be viewed to be identical.

In figure 3, we present the probability density functions $f(\tilde{X})$ of the normalized relative price levels $\tilde{X}$ for both cancelled buy and sell orders of the same four stocks as in figure 2. We find that the PDFs first increase with the normalized relative price level $\tilde{X}$, and then decrease rapidly for both buy and sell orders. Each PDF has its maximum value away from the same best price, which is in line with the results presented in figure 2. In addition, the PDF follows a power-law decay when $\tilde{X} > \tilde{X}_{\min}$,

$$f(\tilde{X} > \tilde{X}_{\min}) \sim \tilde{X}^{-\alpha},$$

where $\alpha$ is the tail exponent of the power law and $\tilde{X}_{\min}$ is the lower threshold of the scaling range of the power-law decay.

There are many methods to determine the power-law tail exponent $\alpha$. Clauset, Shalizi and Newman proposed an efficient quantitative method to estimate $\tilde{X}_{\min}$ and $\alpha$ based on the Kolmogorov–Smirnov test. The Kolmogorov–Smirnov statistic (KS) is defined as

$$KS = \max_{\tilde{X} > \tilde{X}_{\min}} \left( |P - F_{PL}| \right),$$

where $P$ is the cumulative distribution of normalized relative price levels $\tilde{X}$ and $F_{PL}$ is the cumulative distribution of the best power-law fit [18]. One can first determine the threshold $\tilde{X}_{\min}$ by minimizing the KS statistic, and then estimate the power-law tail exponent $\alpha$ with the data in the range $\tilde{X} > \tilde{X}_{\min}$ using the maximum likelihood estimation (MLE) method,

$$\alpha = 1 + m \left[ \sum_{t=1}^{m} \ln \frac{\tilde{X}(t)}{\tilde{X}_{\min}} \right]^{-1},$$

doi:10.1088/1742-5468/2013/04/P04027
The position profiles of order cancellations

Figure 3. Empirical probability density functions $f(\hat{X})$ at log–log scales for both cancelled buy and sell orders of four stocks: (a) 000 012, (b) 000 021, (c) 000 066 and (d) 000 0625. The solid lines are the fitted functions using the maximum likelihood estimation method.

where $m$ is the number of the data points in the range $\hat{X} > \hat{X}_{\text{min}}$. The standard error $\hat{\sigma}$ is derived from the width of the likelihood maximum, which can be expressed as follows

$$\hat{\sigma} = \frac{\alpha - 1}{\sqrt{m}}.$$  (7)

We analyze the PDFs of normalized price levels $\hat{X}$ for other stocks and find that they all exhibit power-law right tails. Applying the method mentioned above, we determine the characteristic parameters $\alpha$, $\hat{X}_{\text{min}}$ and $\hat{\sigma}$, which are illustrated in table 4. We find that the power-law exponent $\alpha$ varies in the range [1.80, 2.51] for buy cancellations and [1.74, 2.42] for sell cancellations. The exponents $\alpha$ are all close to 2, with the mean value $\overline{\alpha} = 2.06 \pm 0.16$ for cancelled buy orders and $\overline{\alpha} = 2.12 \pm 0.17$ for cancelled sell orders. Moreover, for 8 stocks, the tail exponents of the cancelled buy orders are greater than the cancelled sell orders.

Because the PDFs of normalized relative price levels $\hat{X}$ for the 23 stocks are similar to one another, we treat the 23 stocks as an ensemble and determine the PDF with the 23-stock aggregating data. As expected, the ensemble PDF is similar to the individual stock, with a power-law relaxation in the right tail. Using the KS-test and maximum likelihood estimation method, we obtain $\alpha = 2.02$ with $\hat{X}_{\text{min}} = 7.34$ for cancelled buy orders and

doi:10.1088/1742-5468/2013/04/P04027
The position profiles of order cancellations

Table 4. Characteristic parameters of PDFs of normalized relative price levels $\hat{X}$ for both buy and sell cancellations of the 23 stocks. $\alpha$, $\hat{X}_{\text{min}}$ and $\hat{\sigma}$ are the parameters of the power-law tail distribution based on the KS-test and likelihood maximum estimation method.

| Stock | Buy orders | | | Sell orders | | |
|-------|------------|--|---|-------------|--|---|
|       | $\alpha$  | $\hat{\sigma}$ | $\hat{X}_{\text{min}}$ | $\alpha$  | $\hat{\sigma}$ | $\hat{X}_{\text{min}}$ |
| 000 001 | 1.82 | 0.001 | 4.85 | 2.09 | 0.002 | 10.61 |
| 000 002 | 1.96 | 0.005 | 4.07 | 2.16 | 0.006 | 8.54 |
| 000 009 | 2.01 | 0.002 | 4.06 | 2.35 | 0.003 | 10.41 |
| 000 012 | 1.96 | 0.003 | 5.63 | 2.07 | 0.003 | 9.18 |
| 000 016 | 2.07 | 0.004 | 5.04 | 2.19 | 0.005 | 11.10 |
| 000 021 | 2.05 | 0.003 | 11.23 | 2.19 | 0.003 | 9.75 |
| 000 024 | 2.10 | 0.005 | 7.75 | 1.90 | 0.004 | 3.94 |
| 000 027 | 2.08 | 0.006 | 7.73 | 2.06 | 0.005 | 10.88 |
| 000 063 | 2.18 | 0.007 | 18.49 | 2.26 | 0.007 | 19.34 |
| 000 066 | 2.10 | 0.003 | 9.79 | 2.22 | 0.004 | 18.98 |
| 000 088 | 2.51 | 0.026 | 19.39 | 2.42 | 0.028 | 22.00 |
| 000 089 | 2.11 | 0.008 | 6.36 | 1.93 | 0.006 | 3.68 |
| 000 406 | 2.01 | 0.003 | 4.89 | 2.13 | 0.004 | 19.40 |
| 000 429 | 2.10 | 0.006 | 3.30 | 2.19 | 0.006 | 15.17 |
| 000 488 | 2.45 | 0.012 | 15.53 | 1.98 | 0.005 | 17.29 |
| 000 539 | 2.04 | 0.006 | 5.67 | 2.23 | 0.008 | 27.91 |
| 000 541 | 1.95 | 0.007 | 3.79 | 1.79 | 0.006 | 2.76 |
| 000 550 | 1.99 | 0.003 | 6.24 | 2.19 | 0.003 | 13.87 |
| 000 581 | 2.12 | 0.007 | 7.86 | 2.01 | 0.006 | 11.04 |
| 000 625 | 2.00 | 0.003 | 9.31 | 2.18 | 0.003 | 12.15 |
| 000 709 | 2.02 | 0.004 | 3.26 | 2.23 | 0.005 | 12.01 |
| 000 720 | 1.80 | 0.006 | 3.35 | 1.74 | 0.006 | 3.83 |
| 000 778 | 2.07 | 0.005 | 7.55 | 2.19 | 0.005 | 18.74 |

$\alpha = 2.15$ with $\hat{X}_{\text{min}} = 10.59$ for cancelled sell orders, which are close to the mean values ($\overline{\alpha} = 2.06$ for buy cancellations and $\overline{\alpha} = 2.12$ for sell cancellations), respectively.

4. The profile of cancellation positions at the price level

In this section, we study the profile of cancellation positions at a certain price level, and apply the probability density function (PDF) to analyze it as well. Denote $y(x, t)$ as the position of a cancelled order in the queue of the $x$th price level at time $t$. For example, the cancelled order marked with gray color in figure 1 is located at the second position at the second price level ($x = 2$) in the buy LOB at time $t$, i.e. $y(2, t) = 2$. Taking into account the impact of the number of orders stored at the $x$-th price level, we define the relative position variable $Y(x, t)$ instead of $y(x, t)$,

$$Y(x, t) = \frac{y(x, t)}{n_{b,s}(x, t)},$$

(8)
The position profiles of order cancellations

Figure 4. Probability density functions $f(Y)$ of relative positions $Y(x,t)$ at the first price level ($x = 1$) (a), the second price level ($x = 2$) (b), the third price level ($x = 3$) (c), and the fourth price level ($x = 4$) (d) for both buy and sell orders of stock 000 001. The solid curves are the best fitting function illustrated in equation (9).

where $Y(x,t)$ varies in the range $(0, 1]$, and $n_{b,s}(x,t)$ is the total number of orders stored at the $x$-th price level at time $t$ in the buy or sell LOB.

Now we analyze the probability density functions $f(Y)$ of relative positions $Y(x,t)$ at the $x$-th price level for both buy and sell cancellations. The PDFs $f(Y)$ of the first four price levels ($x = 1 - 4$) of stock 000 001 are illustrated in figure 4, and we find the PDFs have similar shapes. Furthermore, the function $f(Y)$ is close to zero when the relative position $Y$ approaches zero. As $Y$ increases, $f(Y)$ first increases rapidly in the range $Y \leq 0.1$, then it fluctuates around a constant level until the end of the LOB. As we can see in the figures, the probability of order cancellation is small in the front of the queue at the first four price levels, which means that orders earlier submitted are less likely to be cancelled. It may be caused by the reason that in the Chinese stock market, most of the traders are patient and FO risk is not the main reason for cancellation.

In order to approximate the PDF of relative positions, we empirically apply an exponential function, which reads

$$f(Y) = \frac{1}{z}(1 - e^{\beta Y}),$$

where $\beta$ is request exponent and $z$ is a normalized constant, that is, $z = \int_0^1 (1 - e^{\beta x}) dx$. 

doi:10.1088/1742-5468/2013/04/P04027

13
The position profiles of order cancellations

Figure 5. Probability density functions $f(Y)$ of aggregating relative positions for both buy and sell orders of four stocks, 000 012 (a), 000 021 (b), 000 066 (c) and 000 625 (d). The solid curves are the best fitting function illustrated in equation (9).

Table 5. Characteristic parameters of PDFs of the first four price levels for both buy and sell orders of stock 000 001. $\beta$ is the exponent of the fit function expressed in equation (9) and $rms$ is the root mean square of the difference between the best fit and the empirical data mentioned in table 3.

| Price level ($x$) | Buy orders | Sell orders |
|-------------------|------------|-------------|
|                   | $\beta$    | $rms$       | $\beta$    | $rms$       |
| 1                  | $-30.34$   | $0.22$      | $-21.51$   | $0.30$      |
| 2                  | $-24.89$   | $0.22$      | $-25.78$   | $0.26$      |
| 3                  | $-22.25$   | $0.21$      | $-29.54$   | $0.24$      |
| 4                  | $-25.82$   | $0.20$      | $-20.35$   | $0.20$      |

Using the least squares fitting method, we estimate the exponents $\beta$ of relative positions $Y(x,t)$ ($x = 1 - 4$) for both buy and sell orders of stock 000 001, which are presented in table 5.

We also analyze the PDFs of relative positions for randomly chosen price levels and find that it has a similar distribution. So we aggregate the data $Y(x,t)$ at all price levels together, and calculate the ensemble PDFs of relative positions. Figure 5 presents the
Table 6. Characteristic parameters of PDFs for both buy and sell orders of 23 stocks. The parameters $\beta$ and $\text{rms}$ have the same meaning as in table 5.

| Stock  | Buy orders |  | Sell orders |  |
|--------|------------|---|-------------|---|
|        | $\beta$    | $\text{rms}$ | $\beta$    | $\text{rms}$ |
| 000001 | -33.78     | 0.17 | -36.57     | 0.16 |
| 000002 | -9.42      | 0.29 | -14.35     | 0.22 |
| 000009 | -14.07     | 0.22 | -18.90     | 0.19 |
| 000012 | -18.81     | 0.31 | -24.12     | 0.27 |
| 000016 | -12.72     | 0.36 | -18.21     | 0.30 |
| 000021 | -21.26     | 0.29 | -30.91     | 0.23 |
| 000024 | -17.67     | 0.41 | -17.23     | 0.32 |
| 000027 | -13.00     | 0.32 | -13.04     | 0.23 |
| 000063 | -44.70     | 0.36 | -33.26     | 0.24 |
| 000066 | -17.81     | 0.32 | -22.36     | 0.24 |
| 000088 | -32.43     | 0.51 | -21.70     | 0.46 |
| 000089 | -10.89     | 0.41 | -14.46     | 0.31 |
| 0000406| -15.12     | 0.29 | -18.94     | 0.24 |
| 0000429| -10.15     | 0.36 | -12.42     | 0.28 |
| 0000488| -23.40     | 0.41 | -12.32     | 0.37 |
| 0000539| -29.23     | 0.70 | -28.82     | 0.31 |
| 0000541| -9.55      | 0.43 | -16.41     | 0.39 |
| 0000550| -35.81     | 0.30 | -25.01     | 0.25 |
| 0000581| -30.32     | 0.39 | -14.85     | 0.33 |
| 0000625| -40.17     | 0.32 | -22.77     | 0.28 |
| 0000709| -13.25     | 0.26 | -12.91     | 0.24 |
| 0000720| -28.33     | 1.10 | -34.40     | 0.86 |
| 0000778| -12.38     | 0.35 | -17.15     | 0.30 |

PDFs $f(Y)$ of four stocks randomly chosen from the 23 stocks. We find that the aggregated PDF is similar to the PDF at a certain price level illustrated in figure 4. In addition, we estimate the exponents $\beta$ for the rest stocks, and the values are illustrated in table 6.

Moreover, an interesting feature which has been observed is that the PDF $f(Y)$ has periodic peaks at $Y = 0.1m$ ($m = 1, 2, \ldots, 10$) for both buy and sell orders. This pattern is rather robust if we use other bins. This periodic strip pattern is also observed in the snapshot of stocks traded on the London Stock Exchange [19], in the LOB shape in the Chinese stock market [11], and in the PDF of order sizes in the opening call auction in the Chinese stock market [20]. The underlying mechanisms of these periodic behaviors might be different, most of which are still unknown. For $f(Y)$, the mechanism is rather mechanical and stems from the discreteness of the length $n_{bs}$ of the queue. The two largest peaks at $Y = 0.5$ and 1 can be easily explained. Assume that any position $y$ has the same cancellation probability for a given queue. For any queue of length $n_{bs}$ in the limit-order book, the occurrence probability of $Y = 1$ is $P(Y = 1) = 1/n_{bs}$. For a queue of even length, we have $P(Y = 0.5) = 1/(2n_{bs})$. When we average the probability over many queues with different lengths, the values of these two $Y$’s stand out.

In the general case of all peaks, the interpretation is not straightforward. We find that the histogram shape of $n_{bs}$ is quite similar to the order book shape [11]. It increases to a
The position profiles of order cancellations

Figure 6. Probability density functions $f(Y)$ obtained from the numerical experiment.

peak around $n_{bs}=10$ and then decays. The relatively large occurrences of $n_{bs}=10$ already imply the possible appearance of peaks at $Y=0.1m$ in $f(Y)$. To test the hypothesis that the discreteness of the queue length causes those peaks, we perform a simple simulation. We assume that $n_{bs}$ is uniformly distributed in $[1, 100]$. For each queue, the cancellation probability at each position is the same. We have generated $10^6$ queues. The resulting $f(Y)$ is illustrated in figure 6. In this trivial numerical experiment, we also observe peaks at $Y=0.1m$ with $m=1, 2, \ldots, 10$. This is direct evidence supporting our explanation.

5. Conclusion

Cancellation plays an important role in the dynamics of price formation in financial markets and helps traders to avoid NE risk or FO risk. In this paper, we have analyzed the order flow data of 23 liquid stocks traded on the Shenzhen Stock Exchange in the year 2003 and studied the empirical the position profiles of order cancellations by rebuilding the buy and sell limit-order books. The main aim of this work is to provide information about order cancellations, which is expected to be used in the construction of zero-intelligence order-driven models [19, 21].

We first analyzed the position profiles of relative price levels where cancellations locate and find that the probability density function (PDF) obeys the log-normal distribution. It implies that when impatient traders cannot make an transaction immediately, they will cancel the placed orders and place new orders to catch the price in order to avoid NE risk. Moreover, we observed that the PDF of cancelled sell orders has a higher peak near the same best price than the buy cancellations. The result is consistent with the fact that the Chinese stock market was bear from 2001 to 2005 and traders were more inclined to submit sell orders. We further studied the PDF of cancellation positions for the normalized relative price levels, and found that the PDF follows a power-law behavior in the tail. We also studied the PDF of cancellation positions at a certain price level, and found that the PDF increases rapidly in the front of the queue and then fluctuates around a constant value until the end of the LOB. In addition, the PDF can be fitted by exponent
functions for both buy and sell cancellations. It means that in the Chinese stock market, orders earlier submitted are less to be cancelled, which may be due to most traders being patient and FO risk not being the main reason for cancellation.

Further work can be conducted to analyze the statistical properties of the lifetime of cancelled orders and to study the relation among cancellation position, order lifetime, order size and so on. It will certainly be interesting.

Acknowledgments

This work was partly supported by National Natural Science Foundation of China (Grants nos 71101052, 71131007 and 11075054), Shanghai ‘Chen Guang’ Project (Grant no. 2010CG32), Shanghai Rising Star (Follow-up) Program (Grant no. 11QH1400800), and the Fundamental Research Funds for the Central Universities.

References

[1] Griffiths M D, Smith B F, Turnbull D A S and White R W, The costs and determinants of order aggressiveness, 2000 J. Financ. Econ. 56 65
[2] Fong K Y L and Liu W M, Limit order revisions, 2010 J. Bank. Finance 34 1873
[3] Yeo W Y, Cancellations of limit orders, 2005 Working Paper National University of Singapore
[4] Coppejans M and Domowitz I, Screen information, trader activity, and bid-ask spreads in a limit order market, 1999 Working Paper Duke University
[5] Liu W M, Monitoring and limit order submission risks, 2009 J. Financ. Mark. 12 107
[6] Hasbrouck J and Saar G, Limit orders and volatility in a hybrid market: the Island ECN, 2002 Working paper New York University
[7] Biais B, Hillion P and Spatt C, An empirical analysis of the limit order book and the order flow in the Paris Bourse, 1995 J. Finance 50 1655
[8] Ellul A, Holden C W, Jain P and Jennings R, Determinants of order choice on the New York Stock Exchange, 2003 Working Paper Indiana University
[9] Ellul A, Holden C W, Jain P and Jennings R, Order dynamics: recent evidence from the NYSE, 2007 J. Empir. Finance 14 636
[10] Ni X-H, Jiang Z-Q, Gu G-F, Ren F, Chen W and Zhou W-X, Scaling and memory in the non-Poisson process of limit order cancellation, 2010 Physica A 389 2751
[11] Gu G-F, Chen W and Zhou W-X, Empirical shape function of limit-order books in the Chinese stock market, 2008 Physica A 387 5182
[12] Bouchaud J-P, Mézard M and Potters M, Statistical properties of stock order books: empirical results and models, 2002 Quant. Finance 2 251
[13] Potters M and Bouchaud J-P, More statistical properties of order books and price impact, 2003 Physica A 324 133
[14] Challet D and Stinchcombe R, Analyzing and modeling 1 + 1d markets, 2001 Physica A 300 285
[15] Maslov S and Mills M, Price fluctuations from the order book perspective—empirical facts and a simple model, 2001 Physica A 299 234
[16] Weber P and Rosenow B, Order book approach to price impact, 2005 Quant. Finance 5 357
[17] Eisler Z, Kertész J and Lillo F, The limit order book on different time scales, 2007 Proc. SPIE 6601 66010G
[18] Clauset A, Shalizi C R and Newman M E J, Power-law distributions in empirical data, 2009 SIAM Rev. 51 661
[19] Mike S and Farmer J D, An empirical behavioral model of liquidity and volatility, 2008 J. Econ. Dyn. Control 32 290
[20] Gu G-F, Ren F, Ni X-H, Chen W and Zhou W-X, Empirical regularities of opening call auction in Chinese stock market, 2010 Physica A 389 278
[21] Gu G-F and Zhou W-X, Emergence of long memory in stock volatilities from a modified Mike-Farmer model, 2009 Europhys. Lett. 86 48002

doi:10.1088/1742-5468/2013/04/P04027