

Research Article

The Price Impact of Order Book Events from a Dimension of Time

Wentao Chi ¹, Xuemei Zhao ², and Lufei Huang ¹

¹*School of Financial Technology, Shanghai Lixin University of Accounting and Finance, Shanghai 201209, China*

²*Shandong Management University, Jinan, Shandong 250357, China*

Correspondence should be addressed to Lufei Huang; lufei.huang@lixin.edu.cn

Received 16 April 2021; Accepted 19 July 2021; Published 27 July 2021

Academic Editor: Xiaobo Qu

Copyright © 2021 Wentao Chi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a new linear model to explain the price move by Level-2 high-frequency data in Chinese mainland stock market. In Chinese stock market, the cancellation ratio is very low, and imbalanced order flow prevails most of the time in the trading periods. From time dimension viewpoint, we find the difference of efficiency of limit orders executed, respectively, in bid/ask limit order book, order execution imbalance (OEI), could improve the classic model of Cont et al. (2014) based on market microstructure of Chinese mainland stock market. In particular, when market's liquidity is booming, our model's explanatory power and *R*-squared increased sharply. And the correlations of OEI are very high that may be exploited to predict the price move in the next time window for doing high-frequency trading.

1. Introduction

Excessive order cancellations are scrutinized by regulators who view such excess as a possible indicator of manipulative quoting activity by potential stock market manipulators. In US stock market, Hasbrouck and Saar [1] investigate trading of 100 NASDAQ-listed equity securities on INET, an electronic limit order book, and find that over 35% of limit orders are cancelled within two seconds of submission compared with only 5%, which we computed from most liquidity stocks in Shenzhen stock exchange in Chinese stock market. The market microstructure from China will contribute very different order flows from US market composed of limit orders, market orders, and cancellation orders, which are usually discussed in high-frequency trading.

The field of market microstructure, in which reasons could cause price moving in continuous trading period, is a research mainstream, because many factors affect price variation like big market orders, clouded market orders on the same side, breaking news, market liquidity, and number of participant members. Until recently, there are no clear factors and models to determine price moving direction and altitude from high-frequency trading data especially in Chinese mainland stock exchange market. First, we test the order flow imbalance (OFI) factor whether or not it can contribute most

R-squared for price moving in every ten seconds from the most liquidity 50 stocks in Shenzhen stock market for one month and compare the results within another month, in which trading volume blooming in recent years. Then, we find that although statistical results derived from OFI are coherent to findings of Cont et al. [2], the *R*-squared is not as high as NYSE's based on the research of Cont et al. [2]. In practical high-frequency trading, we find that analysis of actions on order book from time dimension is critical for HFT especially in the period of intensive trading activity. Like the reverse of limit orders trade duration, or difference between ratios of arrived buy and sell limit executed, respectively, in a short and fixed timespan, order execution imbalance (OEI) is a time dimension reflecting potential market force to move price up/down. And the time dimension factor model based on Level-2 data of Chinese stock market effectively improves the *R*-squared compared with Cont's model, and our theory is coherent to principles of market microstructure. In the end, we show that when total market liquidity is surging, these explanatory power and *R*-squared of our model will be augmented sharply.

Motivated by the above research, we show the price impact model with a time dimension of these orders. The time dimension factor model based on Level-2 data of Chinese stock market effectively improves the *R*-squared

compared with Cont's model, and our theory is coherent to principles of market microstructure. In the end, we show that when total market liquidity is surging, this explanatory power and R -squared of our model will be augmented sharply.

The paper is structured as follows. In Section 2, the related literature is reviewed. In Section 3, we compare Chinese stock market with US stock market with perspective of proportion of balanced/imbalanced order flow and order cancellation ratio. And then, we provide a new method from time dimension viewpoint, which shows that out of synchronization for limit order books or order queues constructed and deleted by market orders in the bid and ask side respectively could lead to price changes, and other features of order book are unchanged. After that, based on Cont's definition about OFI, we specify a linear model to explain price changes about OFI and OEI. In Section 4, we use 50 most liquid stocks from Shenzhen stock exchange of Chinese mainland to estimate the coefficients of variables and analyze what we find in a normal month compared with a trading volume booming month recently, and then we get statistical results for eight different time periods of a trading day. In Section 5, based on the empirical results, we find high correlations of order execution imbalance (OEI) in high-frequency trading to predict price moves before adverse selection, which is caused by price changes when we placed limit orders waiting a long time in the order book. In Section 6, we conclude possible reasons why OFI and OEI still have deficiencies in explaining price changes and finally propose potential methods for further improvement.

2. Related Literature

2.1. Related Work of Macrofinancial Markets. The moving direction and altitude of prices in financial markets result from the interaction of buy and sell orders through a complex dynamic process. The availability of high-frequency records of orders, trades, and quotes has reported statistical regularities in limit order book (LOB) data from a wide variety of different markets. LOBs are subject to frequent shocks in order flow that cause them to display nonstationary behavior, thus, in the result cause price impact. Ellul et al. [3] reported a positive correlation between higher midprice realized volatility and the percentage of arriving orders that were limit orders. The intuition behind price moving is an imbalance between supply and demand order flows. Cont et al. [2] show that, over short time intervals, price changes are mainly driven by the order flow imbalance (OFI), defined as the imbalance between supply and demand at the best bid and ask prices. But the state space of order book is very large conditioning on the fact that the most recent event is still problematic. Findings from Cont et al. [2] seem to give an intuitive picture of the price impact of order book events, which is somewhat simpler than the ones conveyed by previous studies. Meanwhile, Cont's linear model with average high R -squared also excludes trades, which seem to carry little to no information about price changes after the OFI is taken into account simultaneously.

2.2. Related Work of Trading Strategies. Achab et al. [4] introduce a new nonparametric method that allows for a direct, fast, and efficient estimation of the matrix of kernel norms of a multivariate Hawkes process. Dugast [5] studied the same model and proposed a prediction that positive (negative) market order imbalance, negative (positive) depth, and cancellation imbalances contribute a positive (negative) change in price. Following market news, he found that order flows become unbalanced, and market depth is consumed, leading to positive covariance between price variability and order book unbalances. Prior to news arrival, trading occurs because of differences in private valuations, though at prices generally in line with the asset value. Yet when news arrives, trading prices no longer accord with the new asset value. This mismatch generates imbalances, in both order book and order flows, that disappear once prices have adjusted. Huang et al. [6] are interested in whether the combined estimator may be used to form a combined forecast to improve the RE forecast (forecast made using the RE estimator) and the FE forecast (forecast using the FE estimator) in out-of-sample forecasting.

Since imbalances anticipate a change in price, it follows that they could be exploited, especially by algorithmic trading strategies. Cartea et al. [7] document the predictive power of order book imbalances for future price movements on the Nasdaq exchange. Goldstein et al. [8] show that HFTs on the Australian Securities Exchange take advantage of this predictability. Silantsev [9] conducts an in-depth analysis on the trade and quote data of the XBTUSD perpetual contract and demonstrates that the trade flow imbalance is better at explaining contemporaneous price changes than the aggregate order flow imbalance. Niu et al. [10] studied the valuation of vulnerable European options incorporating the reduced-form approach, which models the credit default of the counterparty. Fosset et al. [11] proposed an actionable calibration procedure for general Quadratic Hawkes models of order book events and found that the Zumbach kernel is a power-law of time, as are all other feedback kernels.

2.3. Literature Summary. From Chinese mainland stock market, we found that the order flow imbalance cannot explain much for price change. Here, today, more and more automated or algorithms trading institutions and professional traders enter Chinese stock market to do T-0 trading (close holding positions everyday) with established stock positions one day ago because of Chinese T-1 trading limitations. Most professional T-0 traders submit limit orders with anticipation of their limit orders that could be matched as soon as possible, especially in 30 minutes after market opened or 30 minutes before market closing, in which most market trading volume accumulated in these periods. With orders' time priority rule, time creates new incentives to use trading information extending Cont's model that would be a worthwhile contribution in its own right. Computing orders' execution timespan once limit orders are submitted into Cont's classical linear model contributes more explanation power for price changes in a 10-second interval from Chinese mainland stock market.

Thus, adding the factor of difference between the ratio of sell limit orders and buy limit orders executed in a fixed timespan proportional to total sell/buy limit orders' arriving at order book in every 10 seconds is a time dimension reflecting potential force to move price up/down. And empirical results illustrated that this is effective.

3. Price Impact Model

3.1. New York Stock Exchange and Nasdaq versus Chinese Mainland Stock Exchanges. Every day, 70% trading volume is from high-frequency trading in US exchange market. Technology has changed financial markets, altering the trading behavior of limit order traders. High-speed computerized trading strategies, and electronic order-driven trading exchanges, let traders better monitor their orders and make faster, more accurate decisions. The increase in trading speed coincides with an explosion in order cancellation activity (Hasbrouck and Saar, 2013). Trading in financial markets has entered the nanosecond age in US financial market, where liquidity is added and subtracted in billionths of a second.

In Chinese stock market, the algorithm and high-frequency trading are at most 10 percent in everyday trading volume analyzed by stock exchanges. And the solution of time stamp from limit orders, market orders, or other kinds is correct to 10 milliseconds both in Shenzhen stock exchange and Shanghai stock exchange. This data resolution would be an obstacle for high-frequency traders in Chinese market. And moreover, Chinese SEC and stock exchanges limit orders' cancellation. Table 1 shows that cancellation ratio is as low as 5% or less.

Table 1 shows the estimation for ratios of orders' cancellation (column 3 for bid side and column 5 for sell side) and ratios of order flow imbalance (column 2 for bid side and column 4 for sell side) from Shenzhen stock exchange; $\lambda(b)$, $\mu(b)$, and $\theta(b)$ stand for limit orders arriving rate, market orders arriving rate, and orders' cancelling rate, respectively, in number of shares per second on March, 2019, by averaging 50 stocks. And the same parameters' descriptions are for sell side.

The order cancellation rate is represented with theta θ , limit order arrival rate is represented with λ , and market order arrival rate is represented with μ . Because, in different periods of a trading day, the liquidity and trading activity varies vastly, Table 1 summarizes cancellation ratios (columns 3 and 5) and unbalanced order flow ratios (columns 2 and 4) for eight trading periods, in which the intensities of cancelled orders and executed orders are compared with limit orders for bid side and sell side (ask side), respectively.

In contrast to Chinese financial market, Hasbrouck and Saar [12] investigate trading of 100 NASDAQ-listed equity securities on INET, an electronic limit order book, and find that over 35% of limit orders are cancelled within two seconds of submission.

Cont and de Larrard [13] consider the case of a balanced order flow, for which the arriving intensities of market orders and cancellations are equal to arriving intensity of

limit orders. The study of high-frequency quote data indicates that this is an empirically relevant case for many liquid stocks. For the buy side, they find that the imbalance between arriving intensity of limit orders and intensities of market orders + cancellations is around 5% or less for these stocks. But Table 1 shows that the imbalance is around 33% in Chinese mainland stock market.

So, controlled by Chinese SEC on limit orders' cancellation, the state of imbalanced order flow dominates the market since the proportion of cancellation is low compared with total limit orders in Chinese stock market. The order flow imbalance (OFI), proposed by Cont et al. [2], is defined as the imbalance between supply and demand at the best bid, and ask prices better explain price changes. Their linear model's goodness of fit is surprising for high-frequency data with a R -squared of 65% on average across 50 stocks in S&P 500 constituents.

Hence, in the environment of imbalanced order flow dominated in Chinese stock market, maybe we cannot establish a stable linear relationship or with high R -squared model between OFI and price changes because imbalanced order flow is so prevailing in market whenever price changes. And further OFI is constructed from order book events taking place only at the best bid/ask, and results of Cont et al. [2] show that activity at the top of the order book is the most important factor driving price changes. So, the OFI did not try to explain the latitudes of price changes. In particular, in the morning of market open, some stocks could pump and dump lots of ticks in seconds or in milliseconds in Chinese stock market.

Since the limitation on orders' cancellation, algorithms, and high-frequency trading firms in Chinese mainland hopes that their limit orders can be filled as soon as possible to avoid adverse selection when trading span or duration of limit orders, which is difference between arriving time and first filling time of limit orders, is getting longer, some algorithms trading firms usually use TWAP or VWAP methods to open or cover their positions by optimal placement strategies, which switch between limit orders and market orders. To ensure saving costs from trading stocks, they will use limit orders as much as possible, and with strict limitation on orders' cancellation, they should estimate orders' filling probability more accurately.

3.2. A Time Dimension Viewpoint and Model of the Limit Order Book. Many high-frequency trading mainly uses market making strategy to place limit orders on different layers into order book for capturing the variance of price. They hope that their placed bid limit orders and ask limit orders within a time interval can be executed almost simultaneously to get bid-ask spread for profits. Like TWAP or VWAP of algorithms trading, traders also try to place limit orders when market price has the potential to move towards to the placed prices for saving costs from slippage or using market orders.

In every trade day, the automated or manual high-frequency trading usually happens at the open of stock markets since, in this period, the prices change quickly, and variance

TABLE 1: In eight different time periods of a trading day, bid(b)/ask(s) side imbalance ratio of orders' arrival (column 2/4) and cancelling ratio of orders (column 3/5).

Time period	$(\lambda(b) - \mu(b) - \theta(b))/\lambda(b)$	$\theta(b)/\lambda(b)$	$(\lambda(s) - \mu(s) - \theta(s))/\lambda(s)$	$\theta(s)/\lambda(s)$
9:30–10:00	0.370073879	0.03662	0.48266014	0.03417
10:00–10:30	0.383481915	0.06472	0.43961763	0.04746
10:30–11:00	0.359117126	0.08013	0.43537183	0.05688
11:00–11:30	0.291588197	0.08751	0.44505638	0.05835
13:00–13:30	0.358005535	0.08612	0.43203355	0.05679
13:30–14:00	0.278926993	0.11276	0.48520687	0.06282
14:00–14:30	0.326217539	0.11834	0.43248183	0.06079
14:30–14:57	0.317330298	0.15007	0.41677333	0.05012

is high, which could cover trading fees. Once they place a limit order, the order may have high possibility to be filled; if the filling possibilities computed are different from the real ones, they have to cancel the previous orders to wait the next execution opportunity.

Execution duration of limit orders is defined as

$$\begin{aligned} \text{ED} &= \text{timestamp of first filled} \\ &\quad - \text{timestamp of arriving into market queue.} \end{aligned} \quad (1)$$

We can see from Figure 1 that if some limit orders arrived at the 5th second at both bid and ask queue, most of them are filled around the 10th second on ask side, and on bid side, the limit buy orders are “consumed” relatively gentle, and the curve represented for matching rate is flat between 5 seconds and 15 seconds. If there are no new limit orders arriving on both sides, at time 10, the midprice will pump suddenly since lots of limit ask orders are deleted from ask queues that make the best ask price move upward, even though the total number of filled orders on bid/ask side is the same on average for interval [5 s, 15 s]. The cluster effect for market orders like in environment of Figure 1 is demonstrated by the fact that the majority proportion of orders in the ask queue filled in a short time interval could cause midprice shift towards to the price higher (market buy orders clustering) or lower (market sell orders clustering).

Then, we use order execution imbalance (OEI) to represent the different ratios of new arriving limit orders, respectively, in bid order book and ask order book if they could be filled in a fixed time interval proportional to total arriving ones. We use two time grids $\{T_0, \dots, T_I\}$ and $\{t_0, 0, \dots, t_{I,K}\}$ with time steps $T_i - T_{i-1} = 30 \text{ min}$ and $t_{k,i} - t_{k-1,i} = \Delta t = 10 \text{ s}$, $[t_{k-1,i}, t_{k,i}] \subset [T_{i-1}, T_i]$. Within each long-time interval $[T_{i-1}, T_i]$, we compute 180 price changes and OEIs indexed by k [14]:

$$\text{OEI}_k = \frac{\sum_{n=L(t_{k-1})+1}^{L(t_k)} \mathbf{1}_{\{ED_n \leq 10\}}^s}{L_k^s} - \frac{\sum_{n=L(t_{k-1})+1}^{L(t_k)} \mathbf{1}_{\{ED_n \leq 10\}}^b}{L_k^b}, \quad (2)$$

where OEI_k is the order execution imbalance in the k th interval; L_k^s is the total limit orders arrived in the k th interval for the sell side; L_k^b is the total limit orders arrived in the k th interval for the buy side; $L(t_{k-1}) + 1$ is the first limit order in the k th interval; $L(t_k)$ is the last limit order in the k th interval; and $\mathbf{1}_{\{ED_n \leq 10\}}^s$ is the indicator function for ED_n less than or equal to 10 seconds in sell side, and the same definition for bid side.

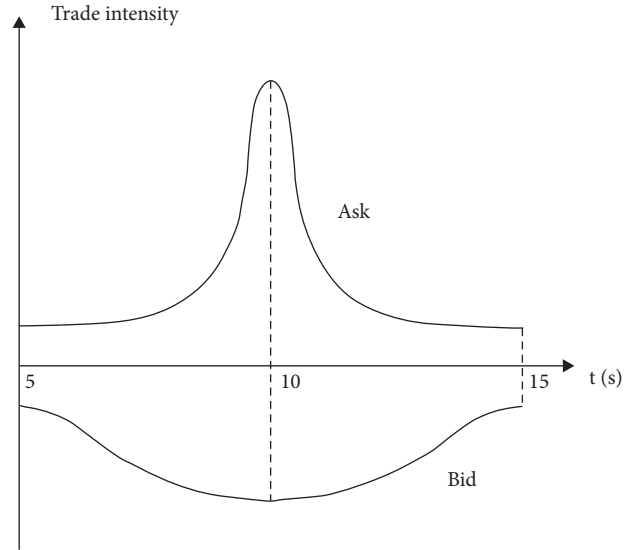


FIGURE 1: Trading intensity evolves differently at bid and ask queue.

Define price changes as in a time interval $[t_{k-1,i}, t_{k,i}]$ the midprice difference between start and end time point.

$$P_{k,i} = \frac{P_N^b(t_{k,i}) + P_N^s(t_{k,i})}{2\delta} - \frac{P_N^b(t_{k-1,i}) + P_N^s(t_{k-1,i})}{2\delta}, \quad (3)$$

where $N(t_{k-1,i})$ and $N(t_{k,i})$ are the index of the first and the last order book events in the interval $[t_{k-1,i}, t_{k,i}]$. The tick size δ is equal to 1 cent in our data. And s stands for sell (ask) side of order book, and b stands for bid (buy) side of order book.

To get OEI, we first add together the number of limit orders arriving at time interval $[t_{k-1,i}, t_{k,i}]$, which will be first filled less than or equal to 10 seconds (execution duration ≤ 10 seconds), and then we compute the ratio of these orders proportional to total limit orders arriving, respectively, for bid and ask side; finally, we get the difference of ratios between sell and buy side. Here, we not only concentrate the orders on best bid/ask queue, but also measure the orders on all different price levels. The OEI reflects which side has more potential willingness to trade, in other words, if part of orders placed by high-frequency traders or algorithms traders are executed in a shorter interval, which provides a sign of whether bid (ask) side is more efficient than the other and has more liquidity for trading than the other. If one of two sides was more efficient for providing “service” higher execution speed for incoming limit orders, then order books

of that side will induce more “customers” limit orders to arrive at this line more frequently. With time elapsing, the denominator L_k^s or L_k^b is getting larger and larger making the “service rate” for further incoming limit orders lower and lower. At last, the queuing system will rebalance to a new state, in which the OEI gets balanced, presented with its absolute value shrinking. Then, prices change contributed by orders’ ED and OEI of order books measured from a time dimension will cease.

Based on the model of Cont et al. [2], the price changes are mainly caused by order flow imbalance:

$$\text{OFI}_k = \frac{L_k^b - C_k^b - M_k^s}{D_b} - \frac{L_k^s - C_k^s - M_k^b}{D_s}, \quad (4)$$

where L_k^b and L_k^s stand for the number of limit orders in the k th interval for bid and ask side, respectively; C_k^b and C_k^s stand for the number of cancellations in the k th interval for bid and ask side, respectively; M_k^s and M_k^b stand for the number of market orders; and D stands for order book depth. In the regression model of Cont et al. [2], they rewrite OFI_k for formula (5) as a variable in empirical findings.

$$\text{OFI}_k = L_k^b - C_k^b - M_k^s - L_k^s + C_k^s + M_k^b. \quad (5)$$

3.3. Model Specification. Here, we establish two linear models: one is the linear model from Cont et al. [2].

$$\Delta P_{k,i} = \beta_i \text{OFI}_{k,i} + \varepsilon_{k,i}, \quad (6)$$

where β_i is a price impact coefficient for an i -th time interval and $\varepsilon_{k,i}$ is a noise term summarizing influences of other factors. It allows β_i and the distribution of $\varepsilon_{k,i}$ to change with index i because of the well-known intraday seasonality effects.

Then, we construct a new linear model including OFI and OEI.

$$\Delta P_{k,i} = \beta_i \text{OFI}_{k,i} + \gamma_i \text{OEI}_{k,i} + \varepsilon_{k,i}. \quad (7)$$

Comparing the statistical results of model (7) with model (6) will show whether or not there is an improvement for model (7) in the environment of high ratio of imbalance order flow and low cancellation rate prevailing in most time of trading day in Chinese mainland stock market.

4. Estimation and Results

4.1. Data. In Chinese mainland, there are two stock exchanges. One is Shanghai stock exchange, and the other is Shenzhen Stock exchange. Totally, around 3000 companies are listed on these two exchanges.

Here, we use the data usually used for high-frequency trading from Shenzhen stock exchange because Shanghai stock exchange does not provide trade by trade files even though there is electronic connection to it from other institutions. So, we lack the trading timestamp of limit orders from Shanghai for computing execution timespan.

Shenzhen stock exchange has three kinds of data to describe what happened in detail when stocks are

trading. The first is the tick file; every three seconds, a new tick will update and disseminate, including tick time with resolution of seconds, newest trade price, best bid price, best bid volume in shares waiting in the queue, best ask price, best ask volume in shares waiting in the queue, and second best bid price. The second is order by order file; every file includes one stock of all orders arriving in the market, including order time stamp with resolution of 10 milliseconds, order price, order volume, index of orders, kinds of orders, and order direction on bid or ask. By the way, all orders cannot be corrected once submitted in Chinese mainland stock market, so if one trader wants to change their previous orders, he/she must cancel them first and then replace orders with new price or volume. The third is trade by trade file; every file includes one stock of all trades happened in the market, including trade timestamp with resolution of 10 milliseconds, trade price, trade volume in shares, label for orders traded or cancelled, index of active orders, and index of passive orders of every trade with the rule of time first and price first; it is a first in first out (FIFO) queuing system, in which every trade happens with a matched pair of one passive order with one active order (usually the market order).

For general, we select 50 stocks with highest liquidity in Shenzhen Stock exchange based on statistics of a month. As shown in Table 2, they summarize 50 most liquidity stocks from Shenzhen stock exchange in July 2018 and March 2019, respectively; in March 2019, the trading volume everyday was booming as high as 150 billion US dollars lasting the whole month.

For comparison, we select data of two months, respectively; one is the month with normal trading volume, and the other is the month with highest trading volume and booming liquid in recent years.

In July 2018, 50 highest liquidity stocks are listed on Shenzhen stock exchange in Table 3 in Appendix, a total of 22 trading days.

We get statistics of stock name, average price in one month, average trading volume in shares for every day, total counts of best bid price updated in one month, total counts of best ask price updated in one month, total counts of trading happened in one month, average bid-ask spread, maximum of spread, and average length of best bid/ask queue in shares for every day.

In March 2019, there are also 50 highest liquid stocks listed on Shenzhen exchange in Table 4 with a total of 21 trading days.

Compared with the two tables, we can see that, in March 2019, even with a total of 21 trading days, the trading frequency and the frequency of best bid/ask price changes are much higher than these in July 2018.

4.2. Empirical Findings. This section reports statistical results of average values across 50 stocks in July 2018 from Shenzhen stock exchange in eight different time periods in a trading day. Model (6) is estimated by an ordinary least squares regression:

TABLE 2: p values, R -squared, and coefficients of $\hat{\gamma}_i$ estimated by model (9) for averaging 50 stocks in March 2019.

Item	9:30–10:00	10:00–10:30	14:30–15:00
p value	0.02157613	0.053005478	0.037117952
R -squared	0.26616222	0.254128669	0.234105648
OEI's coefficient	0.16536644	0.095379223	0.049151509

TABLE 3: Statistical values of highest liquid 20 stocks out of 50 stocks discussed above from Shenzhen stock exchange in July 2018.

Stock name	Price	Volumes	Bidupdates	Iskupdates	Trades	Spread	Max spread	Bidquote	Askquote
"000001"	8.949474	90633922	8584	8585	691911	0.010166	0.05	208784.8	186256.2
"000002"	23.42526	42181755	24764	24490	846306	0.012046	0.17	30181.26	24192.15
"000063"	14.70474	218000000	26186	26151	3877441	0.010325	0.18	7137346	65453.27
"000066"	7.202699	57395362	14472	14450	538278	0.010516	0.07	91451.34	62865.07
"000100"	2.87187	63856220	1674	1670	324921	0.010034	0.02	1231807	1111930
"000333"	47.28303	42269439	43235	42562	1359732	0.014887	0.37	11623.25	9514.582
"000338"	8.516816	62818118	13079	12515	588959	0.010655	0.07	100782.8	95706.92
"000413"	6.029873	53789801	5089	5042	430600	0.010073	0.06	219929.3	199012.8
"000425"	4.045993	69059015	3225	3203	449036	0.010051	0.05	522062.8	460059.8
"000636"	20.15651	94733981	44409	43234	1583531	0.014642	0.41	534288.8	202639.3
"000651"	44.98734	44439886	40025	39113	1510799	0.013861	0.24	12384.81	12172.45
"000709"	2.858653	45590018	954	932	199636	0.010006	0.02	1332800	1342154
"000717"	6.79423	57997266	10213	9987	359260	0.010401	0.09	394507.6	85874.07
"000725"	3.54558	507000000	2397	2389	1970126	0.010006	0.06	4230589	3886116
"000735"	12.02583	111000000	28138	27924	1846659	0.01089	0.18	159588.4	38211.69
"000790"	7.136104	41304190	15006	14632	518929	0.01077	0.1	50848.03	598766.3
"000830"	17.61107	42223402	24782	24311	696922	0.011286	0.1	23583.17	22383.69
"000839"	4.735201	39000627	4820	4731	289843	0.010107	0.05	200474.7	169123.2
"000868"	5.34697	46274097	13327	12987	491337	0.010563	0.12	482389.5	52117.99
"000932"	8.902737	86635903	16526	16407	532922	0.010707	0.08	80327.97	72852.96

TABLE 4: Statistical values of highest liquid 20 stocks out of 50 discussed above stocks from Shenzhen stock exchange in March 2019.

Stock name	Price	Volumes	Bidupdates	Iskupdates	Trades	Spread	Max spread	Bidquote	Askquote
"000001"	12.54927	126000000	17488	17463	1059145	0.01066	0.08	112559.7	108647.9
"000002"	28.63796	74954324	35847	35491	1359969	0.013032	0.15	23236.36	25846.48
"000063"	29.24184	140000000	41597	41023	2894819	0.012885	0.17	1160519	26272.13
"000066"	8.717627	168000000	26081	25436	1249161	0.012592	0.37	3569660	53256.85
"000100"	3.844991	493000000	5356	5305	2378005	0.010047	0.05	8653295	1451183
"000333"	47.81191	35576311	34704	34508	1003728	0.014848	0.25	14357.82	14197.87
"000338"	10.93427	96004436	22551	22309	990645	0.011147	0.1	63859.75	60373.95
"000413"	6.164903	190000000	10079	10029	1246611	0.010096	0.05	283759.9	248469.5
"000425"	4.266872	135000000	5503	5495	744533	0.010089	0.04	457575.5	467730.4
"000636"	15.38505	86504891	30368	29851	1315250	0.011175	0.15	36598.62	31015.7
"000651"	46.03232	55785464	32959	32395	1647012	0.012225	0.21	19192.34	21113.85
"000709"	3.460779	81331901	2176	2085	432321	0.010028	0.02	930606.8	922097.6
"000717"	5.863062	59962873	6285	6181	345876	0.01015	0.09	191616.1	163426.1
"000725"	4.017527	139000000	4420	4405	6170115	0.010009	0.03	4509817	4258081
"000735"	12.91924	101000000	24138	23801	1428595	0.010654	0.11	47920.45	45534.02
"000790"	5.386573	26622979	10210	9866	189760	0.010711	0.13	51580.23	51269.71
"000830"	13.09964	78158583	20499	20213	937589	0.010673	0.1	50154.7	47131.92
"000839"	5.914253	239000000	18407	18086	1492968	0.010551	0.17	2961527	525416.3
"000868"	3.992471	36396440	6204	6076	270629	0.010229	0.25	129570.5	116579.2
"000932"	7.69376	69250545	8615	8395	444436	0.010203	0.05	152788.6	132695.1

$$\Delta P_{k,i} = \hat{\beta}_i \text{OFI}_{k,i} + \hat{\varepsilon}_{k,i}. \quad (8)$$

For ease of show, we get it from the original data to let $\Delta P_{k,i} = \Delta P_{k,i} * \delta$.

Figure 2 presents a scatter plot of ΔP_k against OFI_k of one stock for every 10 seconds in a whole month. There is a

representative stock with an index of SZE.000001, the most liquidity stock in Shenzhen stock exchange.

We can see that the linear effect is not as apparent as the estimations from constituents of S&P 500 by Cont et al. [2]. And the R -squared is relatively much lower compared with stocks in US stock market.

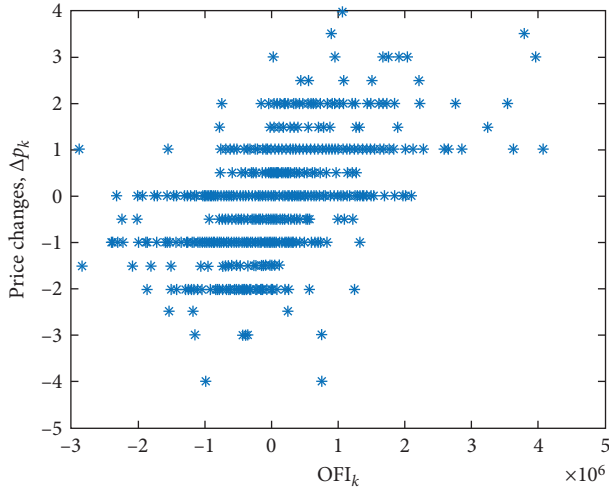


FIGURE 2: Scatter plot of ΔP_k against OFI_k for stock SZE.000001, July 2018.

Table 5 shows that the R -squared is an average of 50 stocks in 22 trading days for 8 different time periods.

The average order book depths for 8 different time periods evolve with time in Figure 3.

And the $\hat{\beta}_i$ of OFI_i changes with time, as shown in Figure 4.

We can see that both $\hat{\beta}_i$ and R -squared of model (8) are decreasing when order book depth gets higher.

And Table 6 shows that p values for coefficient of OFI are all significant at 0.1 threshold. Based on Cont's model, in the second part, we add OEI factor into the linear model to estimate model (7) with improving R -squared efficiently.

Adding values of 8 different time periods together to compute average values of regression model in a trading day is unrealistic. In every different time period, market participants, liquidity, and trading activity are totally different. Usually, high-frequency trading concentrates on open moment of market, at which they were searching trading opportunities that contribute most profits to HFT strategies. Adding the new factor OEI into model (8), we only discuss the statistical results for higher liquidity periods of trading days, such as 9:30 am–10:30 am and 14:30 pm–14:57 pm. In these periods, the most events of order book happened. Then, we estimated the model of

$$\Delta p_{k,i} = \hat{\beta}_i OFI_{k,i} + \hat{\gamma}_i OEI_{k,i} + \hat{\varepsilon}_{k,i}. \quad (9)$$

Table 7 shows that the p values are all significant at 0.1 threshold. And the R -squared increases by 34.3%, 26.8%, and 35.5%, respectively, in model (9) compared to those in model (8).

The curve of Figure 5 for coefficients $\hat{\gamma}_i$ of OEI_i at eight periods is a U-shape showing that when liquidity and trade activities are adequate and intense, the explanatory power of $\hat{\gamma}_i$ is much higher than that in relatively quiet periods of trading.

4.3. Empirical Findings in the Period of Highest Liquidity of Stock Market. The trading volumes and liquidity of stock in Chinese mainland boomed in March 2019, and the statistical results and coefficients of model (8) and (9) change rapidly in this month. We find that the more the liquidity, the

TABLE 5: The R -squared of estimation in model (8).

Time	R -squared
9:30–10:00	0.205776
10:00–10:30	0.190855
10:30–11:00	0.167759
11:00–11:30	0.157325
13:00–13:30	0.156407
13:30–14:00	0.158184
14:00–14:30	0.158095
14:00–14:57	0.144615

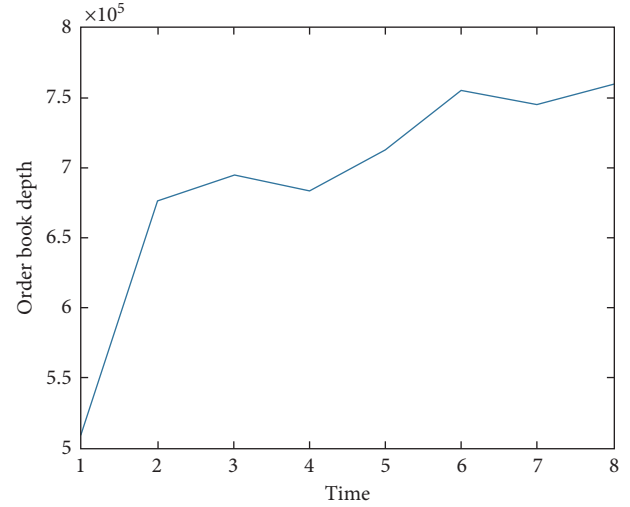


FIGURE 3: The average order book depth at 8 different time periods.

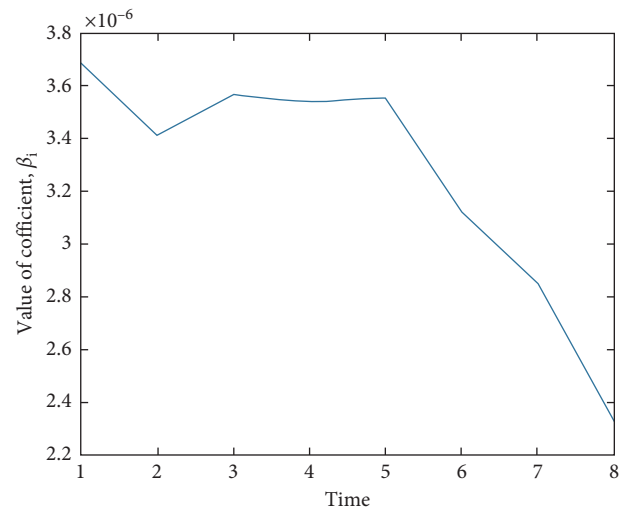


FIGURE 4: The average coefficients of OFI_i for 50 stocks at 8 different time periods.

smaller the p values of OEI_i . The R -squared of model (8) is nearly the same as the R -squared in July 2018. But the R -squared of model (9) and coefficients of $OEI_{k,i}$ increase sharply compared with previous ones in July 2018. Table 8 shows R -squared of model (8) of Cont et al. [2].

TABLE 6: p values for coefficients of OFI_i from model (8) for 8 different trading periods.

No.	p values for coefficient of OFI
1	0.0041350
2	0.0092107
3	0.0160284
4	0.0157925
5	0.0170173
6	0.0181213
7	0.0168114
8	0.0175921

TABLE 7: The R -squared for factor OEI_i and p values for $\hat{\gamma}_i$ in model (9).

Item	9:30–10:00	10:00–10:30	14:30–15:00
p value	0.05385032	0.082507492	0.06868157
R -squared	0.27636538	0.241916241	0.19599882

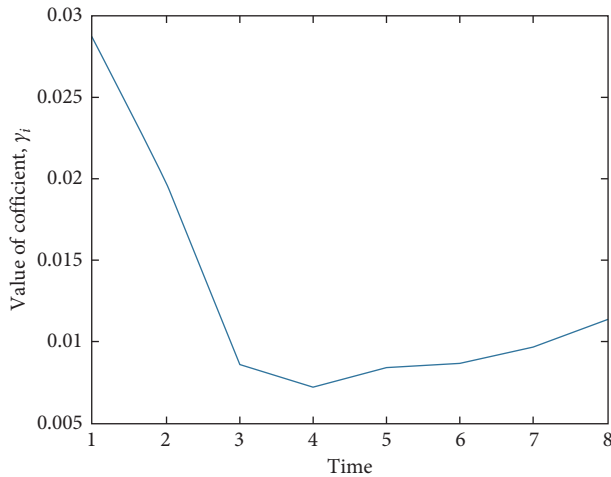


FIGURE 5: The line for values of $\hat{\gamma}_i$ changes along with time.

TABLE 8: The R -squared of model (8) estimated in March 2019.

Time	R -squared
9:30–10:00	0.180138511
10:00–10:30	0.185777441
10:30–11:00	0.175824568
11:00–11:30	0.171979626
13:00–13:30	0.173095659
13:30–14:00	0.173196812
14:00–14:30	0.177265851
14:00–14:57	0.160464719

The values are nearly the same as those in July 2018.

And Table 2 shows the R -squared, p values, and coefficient $\hat{\gamma}_i$ of the factor $OEI_{k,i}$ in model (9), respectively.

From Table 2, we can see that the R -squared is increased by 47.8%, 36.8%, and 45.9%, respectively, in the three actively trading time periods compared with these values in July 2018. And the explanatory power of OEI_i is increased by

477.2%, 387.6%, and 333.9%, respectively, compared with the values (Figure 5) in July 2018.

4.3.1. Summary

- (1) The coefficients of OEI are much higher in actively trading time periods such as the very open moment of market or near closing time of market.
- (2) Within a whole month, the liquidity was booming in March 2019; then the R -squared of our model (9) is enhanced much by near a half of previous values. And more importantly, the coefficients of OEI , explanatory power, are much more than the previous ones.

5. Application of High-Frequency Trading by Prediction of OEI

Predicting price changes in advance is a direct method for trading, but it is too hard to predict if lots of market environment and order book dynamics are included. Whether or not the prediction could be with high rate of winning, market conditions between $[t_{k-1,i}, t_{k,i}]$ and $[t_{k,i}, t_{k+1,i}]$ must have a stable relationship or interplay with each other. And the higher the trading frequency and monitoring frequency for signals, the easier to predict next time direction and momentum of price changes.

In the above discussion of price changes and OFI , we find that price changes in every 10 seconds interval, the autocorrelations are too low to make predictions in next interval, and it is the same situation for OFI even if their autocorrelations are higher compared with price changes in the same period.

But the autocorrelations of OEI are surprisingly high, which reflects that when limit orders waiting on one side could be executed in shorter duration than the other side, the same trend can hold for a while or continue causing the market orders cluster in this side. Thus, we could use the OEI as an indicative of whether or not market orders come to one side order book frequently and densely to form cluster effect in the present and near future moments.

There are high autocorrelations of OEI like Figure 6 for stock SZE.000100 in every 10-second interval.

But autocorrelations of price changes for stock SZE.000100 are nearly ignorable. Figure 7 shows that all the levels of autocorrelations for price changes are very low.

And OFI 's autocorrelations still cannot provide enough evidence for there is strong relationship between OFI in the previous time and the OFI in next time as shown in Figure 8. And the autocorrelations in Figure 8 are not only small but also swinging.

For doing high-frequency trading, switch at the right time between market orders and limit orders before the market price changes, which could cause adverse selection to HFT traders, which is what orders management strategies or optimal placement is doing in real time. How to set and monitor threshold values of market conditions such as OFI , OEI , or other high-frequency indicatives along with peculiarities of the microstructure of Chinese mainland stock market is a future work.

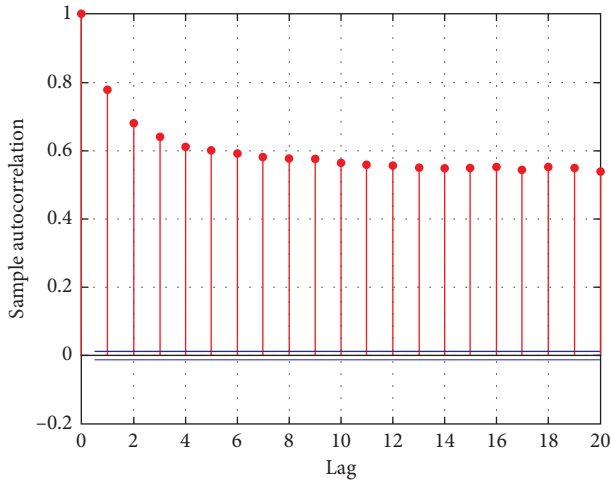


FIGURE 6: Autocorrelations of OEIs for stock SZE.000100 in March 2019.

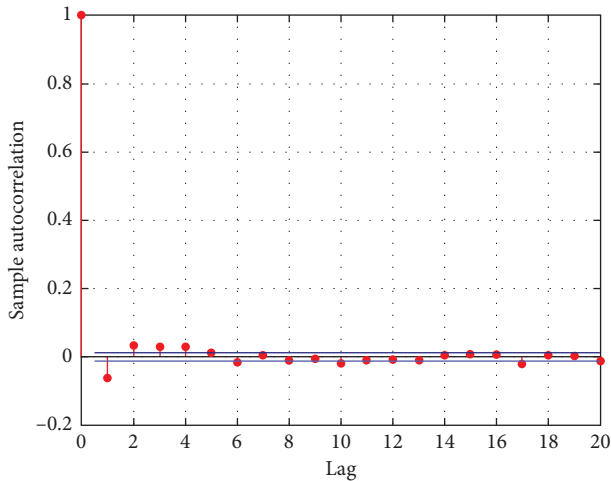


FIGURE 7: Autocorrelations of price changes for stock SZE.000100 in March 2019.

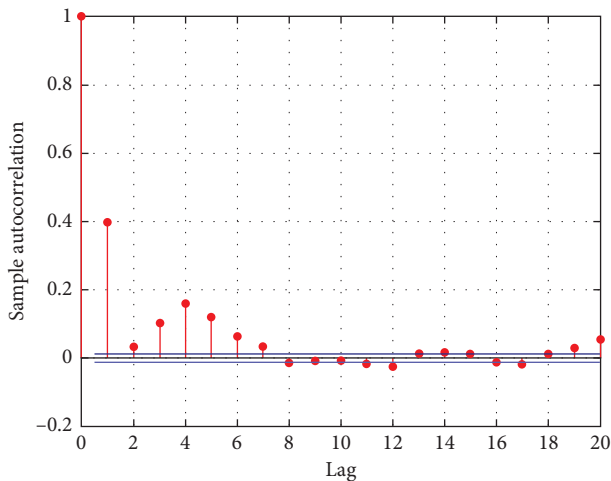


FIGURE 8: Autocorrelations of OFI for stock SZE.00100 in March 2019.

6. Conclusion

In essence, to predict or prescribe price changes is to study the conditional behavior, which is a problem in which the state space of limit order book is huge, and future order flow changes so quickly. Furthermore, in some situations, there is long memory in order flow, and in other situations, they are nearly independent. Therefore, a key modelling task is to find a way to simplify the evolving, high-dimensional state space, while retaining LOB’s key features.

Analyzing and predicting the price changes in above discussions are nonstationary since the R -squared of our model is still not high. Conditioning on most recent events, such as in a 10-second moving window, is problematic. In practice, the most recent events recorded by the exchange may not be the most recent events given a trader observing these data via the trading platform. And in theory, usually, we build a model based on averaging computational results from market conditions for a fix interval but ignoring the prices jump or dump at a particular time point or missing particular distribution information of orders’ properties through taking them averaging into uniform values in a time interval. The missing information such as instantaneous volatility or volatility clustering effect should only be observed from event by event order books. In future work, we could use other methods such as Hawkes process to estimate and analyze the clustering and interplay effects between different orders, which could reflect other conditional information.

Observations of price changes from our model give a brand new viewpoint to capture features of order book dynamics from time dimension of orders’ execution. Not only the R -squared of Cont’s model is enhanced based on Chinese developing stock market, but also, to show it in a different stock market, statistical results gotten from developed stock market uncertainly give us accurate estimations about how the prices, order books, and volatility evolve because some unique features about market microstructure exist in different stock exchanges or markets, although they have the same order-driven matching mechanism.

There are still some limitations in multimethod integration. Many areas that can be further expanded in this study; for example, sustainable development [15], risk interactions [16], multifaceted dimension [17], and innovation network [18] are also the direction of future research.

Data Availability

Data sharing is not applicable to this article as no data sets were generated or analyzed during the current study.

Conflicts of Interest

The authors declare that no conflicts of interest exist regarding this publication.

Acknowledgments

This work was supported in part by the 2020 Shanghai University Key Undergraduate Education Reform Project

“Exploration and Practice of the Construction of New Fintech Majors from the Perspective of the Development of Fintech Industry”.

References

- [1] H. Joel and G. Saar, “Low-latency trading,” *Journal of Financial Markets*, vol. 16, no. 4, pp. 646–679, 2013.
- [2] R. Cont, A. Kukanov, and S. Stoikov, “The price impact of order book events,” *Journal of Financial Econometrics*, vol. 12, no. 1, pp. 47–88, 2014.
- [3] A. Ellul, C. W. Holden, P. Jain, and R. Jennings, “Order dynamics: recent evidence from the NYSE,” *Journal of Empirical Finance*, vol. 14, no. 5, pp. 636–661, 2007.
- [4] M. Achab, E. Bacry, J. F. Muzy, and M. Rambaldi, “Analysis of order book flows using a non-parametric estimation of the branching ratio matrix,” *Quantitative Finance*, vol. 18, no. 2, pp. 199–212, 2018.
- [5] J. Dugast, “Unscheduled news and market dynamics,” *The Journal of Finance*, vol. 73, no. 6, pp. 2537–2586, 2018.
- [6] B. Huang, T.-H. Lee, and A. Ullah, “A combined random effect and fixed effect forecast for panel data models,” *Journal of Management Science and Engineering*, vol. 4, no. 1, pp. 28–44, 2019.
- [7] Á. Cartea, R. Donnelly, and S. Jaimungal, “Enhancing trading strategies with order book signals,” *Applied Mathematical Finance*, vol. 25, no. 1, pp. 1–35, 2018.
- [8] M. A. Goldstein, A. Kwan, and R. Philip, “High-frequency trading strategies,” *SSRN Electronic Journal*, 2018.
- [9] E. Silantyev, “Order flow analysis of cryptocurrency markets,” *Digital Finance*, vol. 1, no. 1, pp. 191–218, 2019.
- [10] H. Niu, Y. Xing, and Y. Zhao, “Pricing vulnerable European options with dynamic correlation between market risk and credit risk,” *Journal of Management Science and Engineering*, vol. 5, no. 2, pp. 125–145, 2020.
- [11] A. Fosset, J.-P. Bouchaud, and M. Benzaquen, “Non-parametric estimation of quadratic Hawkes processes for order book events,” *The European Journal of Finance*, pp. 1–16, 2021.
- [12] J. Hasbrouck and G. Saar, “Technology and liquidity provision: the blurring of traditional definitions,” *Journal of Financial Markets*, vol. 12, no. 2, pp. 143–172, 2009.
- [13] R. Cont and A. De Larrard, “Price dynamics in a Markovian limit order market,” *SIAM Journal on Financial Mathematics*, vol. 4, no. 1, pp. 1–25, 2013.
- [14] K. Chan and W.-M. Fong, “Trade size, order imbalance, and the volatility-volume relation,” *Journal of Financial Economics*, vol. 57, no. 2, pp. 247–273, 2000.
- [15] L. Zhen, L. Huang, and W. Wang, “Green and sustainable closed-loop supply chain network design under uncertainty,” *Journal of Cleaner Production*, vol. 227, pp. 1195–1209, 2019.
- [16] C. Bao, D. Wu, and J. Li, “Measuring systemic importance of banks considering risk interactions: an ANOVA-like decomposition method,” *Journal of Management Science and Engineering*, vol. 5, no. 1, pp. 23–42, 2020.
- [17] A. Díaz and A. Escribano, “Measuring the multi-faceted dimension of liquidity in financial markets: a literature review,” *Research in International Business and Finance*, vol. 51, Article ID 101079, 2020.
- [18] L. Huang, Y. Xu, X. Pan, and T. Zhang, “Green technology collaboration network analysis of China’s transportation sector: a patent-based analysis,” *Scientific Programming*, vol. 2021, Article ID 9961071, 12 pages, 2021.