MA668: Algorithmic and High Frequency Trading
Lecture 19

Prof. Siddhartha Pratim Chakrabarty
Department of Mathematics
Indian Institute of Technology Guwahati

Walking the LOB and Permanent Price Impact

1. We have seen that one of the key ingredients in trading algorithms is how the investor's own actions, together with the order flow of the other market participants affect the price of the assets she/he is trading in.
2. In the trading algorithms developed in latter part of the course, it is shown how strategies depend on the market impact of trades.
3. Instance 1: It is shown how to trade large positions when the investor's own trades walk the LOB, in addition to adversely affecting the mid-price by exerting upward (downward) pressure in the drift of the mid-price if the investor is buying (selling).
4. Instance 2: The problem studied is of an agent wishing to liquidate a large position when the order flow from other traders in the market also impacts the mid-price. In this case, if the agent's execution programme is going with or against the net order flow, the strategy adapts to maximize the revenues from liquidating the position.
5. We want to empirically assess the parameter values for the different effects a trade can have on prices: The permanent and the temporary price impact.

Walking the LOB and Permanent Price Impact (Contd ...)

1. We look at these impacts for five stocks using data from NASDAQ and for the year 2013.
2. A first approach is to estimate these separately.
3. We first estimate the permanent price impact by looking at the impact of order flow on the change in price over five-minute intervals.
4. Let $\overset{\text{Sn}}{\Delta_n} := S_{n\tau} - S_{(n-1)\tau}$ be the change in the mid-price during the time interval $[(n-1)\tau, n\tau]$, where $\tau = 5$ min.
5. Let μ_n be the net order flow defined as the difference between the volumes of buy and sell MOs during the same time interval.
6. We then estimate the permanent price impact as the parameter b in the following robust linear regression:

$$\Delta S_n = b\mu_n + \epsilon_n, \tag{1}$$

   where ϵ_n is the error term (assumed normal).

7. The model (1) is estimated every day, using Winsorised data, excluding the upper and lower 0.5% tails.

Table 4.13

	FARO	SMH	NTAP	ORCL	INTC
$\hat{b}$	1.41×10^{-4}	5.45×10^{-6}	5.93×10^{-6}	1.82×10^{-6}	6.15×10^{-7}
	(9.61×10^{-5})	(4.20×10^{-6})	(2.31×10^{-6})	(7.19×10^{-7})	(2.16×10^{-7})
$\hat{k}$	1.86×10^{-4}	8.49×10^{-7}	3.09×10^{-6}	8.23×10^{-7}	2.50×10^{-7}
	(2.56×10^{-4})	(8.22×10^{-7})	(1.75×10^{-6})	(3.78×10^{-7})	(1.25×10^{-7})
$\widehat{b/k}$	1.02	7.43	2.04	2.28	2.55
	(0.83)	(6.24)	(0.77)	(0.74)	(0.70)

Table 4.13 Permanent and temporary price impact parameters for NASDAQ stocks for 2013. Below each parameter estimate we show its standard deviation.

Figure: Table 4.13

Table 4.13 (Contd ...)

1. The first row shows the average value of the estimated parameters for permanent price impact and the second row shows its standard deviation.
2. The third and fourth rows of the table we show the parameter estimate for temporary impact (denoted by k) and its standard deviation, respectively.
3. To estimate this parameter k, we assume that temporary price impact is linear in the volume traded.
4. Specifically, the difference between the execution price that the investor receives and the best quote is kQ, where Q is the total volume traded.
5. To perform the estimation:
    - A. We take a snapshot of the LOB each second.
    - B. Determine the price per share $S_t^{\text{exec}}(Q_i)$ for various volumes $\{Q_1, Q_2, \ldots, Q_N\}$ (by walking the LOB).
    - C. Compute the difference between the execution price per share and the best quote at that time and perform a linear regression.

Table 4.13 (Contd ...)

1. We regress:

$$\begin{aligned} S_{i,t}^{\text{exec,bid}} &= S_t^{\text{bid}} - k^{\text{bid}} Q_i + \epsilon_{i,t}^{\text{bid}}, \\ S_{i,t}^{\text{exec,ask}} &= S_t^{\text{ask}} + k^{\text{ask}} Q_i + \epsilon_{i,t}^{\text{ask}}, \end{aligned}$$

   where $\epsilon_{i,t}$ represents the estimation error of the i-th volume for the t-th timestamp.

2. The slope argument of the linear regression, $\widehat{k}$ is an estimate of the temporary price impact per share at that time.

3. We do this for every second of every trading day and in the table we report the mean and standard deviation of these daily estimates (for the buy side) when we exclude the first and last half-hour of the trading day and Winsorise the data.

4. Moreover, the fifth row shows the mean of the daily ratio $\widehat{b/k}$ and the sixth row shows its standard deviation.

5. We observe that FARO shows the smallest ratio of 1.02 and SMH shows the largest at 7.43 [a].

---

[a] Later we discuss this ratio in more detail

Figure 4.10

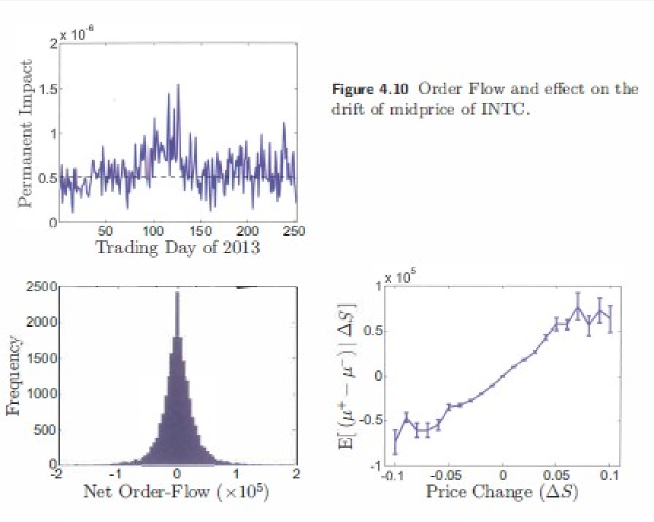

Figure 4.10 Order Flow and effect on the drift of midprice of INTC.

Figure: Figure 4.10

Figure 4.10 (Contd ...)

1. Figure 4.10:
   - A. To showcase the variability of the permanent price impact parameter, the first panel depicts the estimate of b for each day of 2013, with the dashed line showing the average $\widehat{b}$.
   - B. The second panel in the figure shows a histogram of the five-minute net order flow using all the data in 2013.
   - C. Finally, the last panel shows the expected net order flow (with error bars) conditional on a given price change being observed.
2. As already shown by the regression results, there is a positive relationship between net order flow and price changes.
3. The figure shows further details of this relationship to support the finding that when net order flow is positive (negative), that is, more (less) buy than sell MOs, then the mid-price tends to increase (decrease).
4. We see that assuming a linear relationship between price changes and net order flow is plausible for a wide range of mid-price changes.
5. Only in the two extremes, very high or very low price changes, does the relationship fails to be linear, but we note that there are fewer observations in the tails as shown by the histogram.

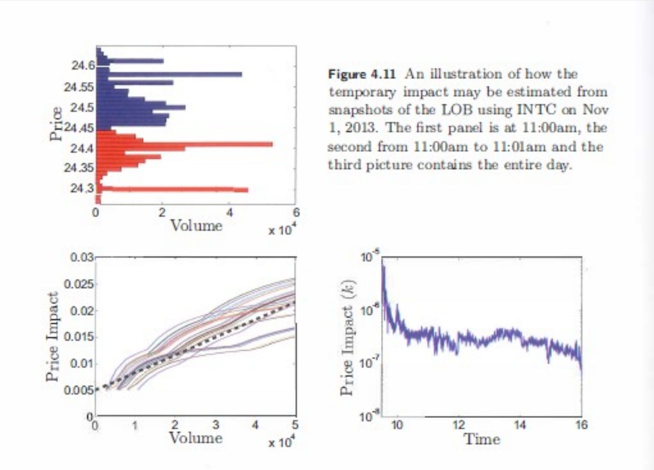

Figure 4.11 An illustration of how the temporary impact may be estimated from snapshots of the LOB using INTC on Nov 1, 2013. The first panel is at 11:00am, the second from 11:00am to 11:01am and the third picture contains the entire day.

Figure: Figure 4.11

Figure 4.11 (Contd ...)

1. Figure 4.11: Explores the temporary price impact for INTC. The top panel shows a snapshot of the LOB for INTC on Nov 1, 2013 at 11am.
2. The bottom left panel captures the empirical temporary price impact curve, generated by hypothetical MOs of various quantities, as they walk through the buy side of the LOB.
3. Each curve represents the curve at every second from 11:00 to 11:01.
4. We also include a linear regression with the intercept set to the half-spread (the dashed line) which would correspond to the model used to estimate the parameter k, earlier.
5. Notice that the impact function fluctuates within the minute and (along) with it the impact that trades of different size could have. The linear regression provides an approximation of the temporary impact during that one minute.

Figure 4.11 (Contd ...)

1. The third picture in the figure shows how the slope of this linear impact model fluctuates throughout the entire day.
2. We see that the largest impact tends to occur in the morning, then this impact flattens and stays flat throughout the day and towards the end of the day it lessens.
3. Such a pattern is seen in a number of assets and is consistent with the reduction in spreads and increases in depth we have documented earlier.
4. The analysis above looks at temporary and permanent effects separately but their joint dynamics is a relevant quantity, in execution algorithms.
5. Liquidation and acquisition strategies take into account the trade-off between costs that stem from walking the book and the permanent impact.
6. In particular, when both types of impact are linear in the rates of trading, this trade-off, is in part captured by the ratio b/k.

Figure 4.12

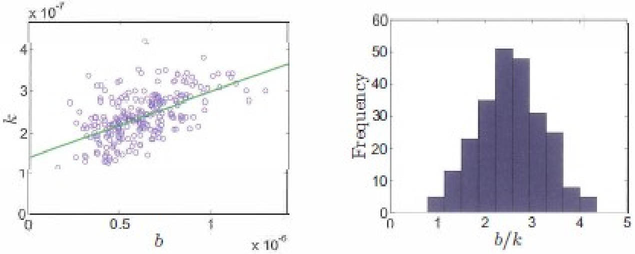

Figure 4.12 Price Impact INTC using daily observations for 2013.

Figure: Figure 4.12

Figure 4.12 (Contd ...)

1. Figure 4.12: Left panel shows a scatter plot of the daily pair (b, k) for INTC which shows a clear positive relationship between temporary and permanent impact.
2. This is consistent with the theoretical relationship between price impact and depth, so that days with little depth will be associated with high price impact, both permanent and temporary, while a deep market will be associated with lower price impacts.
3. Finally, in the right panel of the figure, we see the histogram for the ratio b/k which ranges between 1 and 4 and is symmetric around approximately 2.5.

Messages and Cancellation Activity

1. An important feature in the way exchanges operate is the ability to cancel LOs which have not been filled.
2. Traders who provide liquidity must be able to change their views on the market and therefore cancel their LOs or reposition them in the light of new information.
3. Later: For algorithms that provide liquidity to the market, we will see that these rely on the ability to reposition LOs in the LOB.
4. Example: When we develop market making algorithms that require low latency, the agent is constantly canceling LOs to reposition them, as new information arrives and the agent's view on short-lived deviations in the mid-price are taken into account.
5. Here we employ our detailed ITCH data to measure trading activity by the number of messages recorded by the exchange, where a "message" is a line of data in the ITCH dataset, as we saw earlier.
6. The total number of messages is slightly greater than twice the number of posted orders, as most posted orders are either canceled or executed in full.

Table 4.14

Asset	Mean	StdDev	P01	Q1	Median	Q3	P99
ISNS	1,711	6,078	173	450	760	1,745	8,943
FARO	24,038	10,871	8,524	16,277	22,347	29,232	71,445
MENT	59,661	21,755	23,157	43,477	53,972	72,639	131,715
AAPL	531,728	166,652	280,242	417,576	500,680	614,437	1,067,248

Table 4.14 Daily Number of Messages (in 000s).

Figure: Table 4.14

Table 4.15

Asset	Mean	StdDev	P01	Q1	Median	Q3	P99
ISNS	226.7	749.1	10.8	44.4	80.7	159.1	2885.1
FARO	88.2	55.8	20.7	57.8	79.4	106.1	223.5
MENT	70.0	21.8	29.8	54.2	66.5	83.2	134.2
AAPL	22.6	4.9	12.6	19.3	22.3	25.3	39.4

Table 4.15 Messages per Number of Trades.

Figure: Table 4.15

Table 4.16

Asset	Mean	StdDev	P01	Q1	Median	Q3	P99
ISNS	45.8	3.2	36.3	44.2	46.4	48.3	49.9
FARO	48.1	1.0	44.4	47.6	48.3	48.7	49.5
MENT	47.2	1.0	44.1	46.7	47.4	48.0	48.9
AAPL	43.1	1.9	37.8	41.8	43.3	44.3	47.1

Table 4.16 Cancellations as percentage of Messages.

Figure: Table 4.16

Table 4.14, 4.15 and 4.16

1. Table 4.14: Contains the descriptive statistics for daily messages for our four assets (in thousands).
2. We can see how, as with trading activity, the number of messages for each asset is different by orders of magnitude (except that for MENT which is about 2.5 times that for FARO).
3. In order to adjust for trading activity, a usual procedure is to normalize by the number of trades (as we do in Table 4.15) or by trading volume.
4. Table 4.15: The results suggest an interesting phenomenon: More frequently traded assets "require" fewer messages per trade than less frequently traded ones.
5. Table 4.16: The reverse of this phenomenon is captured, where we look at the percentage of cancellations (out of all messages: posts, cancels and executions). Only for AAPL do we see less than 45% of orders being canceled most of the time.