MA668: Algorithmic and High Frequency Trading
Lecture 13

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

Market Fragmentation

1. Market fragmentation is another issue to be considered while executing aggressively.
2. While this aspect is crucial from the perspective of high-frequency traders, we discuss the matter in brief.
3. Our discussion so far has been focused on detailed data from only one exchange, namely, NASDAQ.
4. October 2014: In the US there are 11 exchanges and around 45 alternative trading venues, most of which are dark pools.
5. Therefore, we cannot truly talk about "the market" as a single exchange, but as the aggregation of activity across a large number of venues.
6. Illustration: To observe market fragmentation, we look at trading during market hours (09:30 to 16:00) for AAPL on July 30, 2013, across all venues using Consolidate Tape data [a].
7. These data provide information on all transactions and best quote from all venues.

---

[a] Wikipedia: Consolidated Tape is an electronic system that collates real-time exchange-listed data, such as price and volume, and disseminates it to investors

Table 3.8

	Percentage Time at NBBO	
Exchange	Bid	Ask
NASDAQ	67.8	61.3
BATS	18.8	15.7
ARCA-NYSE	43.4	38.3
NSE	0.0	0.0
FINRA	0.0	0.0
CSE	0.0	0.0
CBOE	1.2	0.7
EDGA	0.0	0.0
EDGX	34.5	41.0
NASDAQ-BX	0.0	0.0
NASDAQ-PSX	0.0	0.0
BATS-Y	4.5	0.0

Table 3.8 Percentage of Time that Exchange's Best prices are at the NBBO.

Figure: Table 3.8

Market Fragmentation (Contd ...)

1. Table 3.8 captures the information about the percentage of time each exchange's best price (the bid or the ask) coincides with the best price across all venues.
2. Optimally executing a trade is not just about timing and prices displayed in one exchange, but also about:
   - A. How to organize the way an order (or orders) reaches a particular venue.
   - B. Different laws governing how exchanges should handle orders.
   - C. The rules exchanges use to implement them.
   - D. How to programme the routing of order.
   - E. Which order types are best suited to one's particular routing and trading strategy.
3. In the US, the specific regulation Reg. NMS (National Market System) has been set up to facilitate competition between exchanges and to protect investors.
4. In particular, it has specific provisions to protect investors' orders by preventing trade-throughs (execution of an order at an inferior price when a better price is available).

Table 3.9

Exchange	Local Bid	Local Ask	Local Total	NBBO Bid	NBBO Ask	NBBO Total
NASDAQ	22,214	8,122	30,336	458,994	367,181	824,675
BATS	1,854	2,200	4,054	118,205	108,407	225,512
ARCA-NYSE	9,840	5,630	15,470	292,933	273,729	566,361
NSE	901	200	1,101	13,244	11,057	24,301
FINRA	0	0	0	534,178	406,346	940,424
CSE	0	0	0	0	0	0
CBOE	0	0	0	5,005	1,550	6,555
EDGA	0	100	100	31,357	22,125	53,482
EDGX	9,324	2,300	11,624	230,187	207,005	436,392
NASDAQ-BX	1,016	1,100	2,116	60,971	48,365	109,336
NASDAQ-PSX	0	100	100	600	1,525	2,125
BATS-Y	100	1,000	1,100	16,519	16,178	32,497

Table 3.9 Number of Shares Executed at Best Prices (Local refers to best price at local exchange if local best price is not NBBO).

Figure: Table 3.9

Table 3.10

Exchange	NBBO	Local Best	Inside Best	Outside Best	Shares (total)
NASDAQ	39.8	1.5	57.5	1.3	2,073,946
BATS	34.8	0.6	64.2	0.4	648,137
ARCA-NYSE	43.3	1.2	54.4	1.2	1,308,771
NSE	33.1	1.5	64.3	1.1	73,486
FINRA	22.2	0.0	71.5	6.4	4,238,247
CSE	0.0	0.0	0.0	100.0	122,250
CBOE	63.3	0.0	36.7	0.0	10,350
EDGA	27.8	0.1	71.6	0.5	192,202
EDGX	30.6	0.8	67.4	1.1	1,425,145
NASDAQ-BX	41.1	0.8	55.5	3.4	266,166
NASDAQ-PSX	52.1	2.5	40.5	4.9	4,075
BATS-Y	26.0	0.9	73.2	0.0	125,220

Table 3.10 Percentage of Shares Executed, by Execution Quality.

Figure: Table 3.10

Tables 3.9 and 3.10

1. Tables 3.9 and 3.10 look at trade executions in the different venues and compare the price at which the trade was executed "RELATIVE" to the best bid/ask price in the exchange in which the trade is reported (Local) and "RELATIVE" to the best available price across all venues (NBBO: National Best Bid and Offer).
2. Table 3.9: Compares, for each venue, the executions that occurred at the NBBO "VS" the ones at the best price in that venue, when the venue's best price was not the best across all venues.
3. Table 3.10: Allows for further types of executions, not just against a best price, but also inside the (Local) spread (between the local best bid and ask, while not at the NBBO) and outside the spread (for FINRA and CSE we used NBBO as reference, since no local bid/ask quotes are reported).

Empirics of Pairs Trading

1. Most traders do not look at one asset at a time, but consider their interactions between different assets.
2. This works best with groups of assets that share common shocks and occurs naturally for assets in the same industry.
3. Example Considered: Focus on the interaction between a technology stock, Intel (INTC) and a technology ETF, Merrill Lynch Semiconductor ETF (SMH), on November 1, 2013.
4. These two assets move together for two main reasons:
   - A. Mechanical: Around 20% of the ETF holdings are shares in INTC.
   - B. Economic: The ETF is designed to represent the semiconductor industry.

Empirics of Pairs Trading (Contd ...)

1. The analysis of the pairs trading is based on the following theoretical model.
2. We assume that both INTC and SMH have a dynamics given by a mean-reverting and a Brownian motion component, which is represented in the vector form as:

$$d\mathbf{S}_t = \kappa\,(\theta - \mathbf{S}_t)\,dt + \sigma\,d\mathbf{W}_t. \tag{1}$$

3. The presence of a mean-reverting component (the term proportional to dt) introduces the opportunity for generating positive expected returns from trading by exploiting that component's predictability.
4. In this case, we use the joint information from the two processes to create a stronger trading signal by constructing a linear combination of the two assets, which is most strongly driven by the mean-reverting component.

Empirics of Pairs Trading (Contd ...)

1. This is done by transforming the system in (1), which has a matrix κ, into an equivalent system:

$$d\widetilde{\mathbf{S}}_t = \widetilde{\kappa}\left(\widetilde{\theta} - \widetilde{\mathbf{S}}_t\right) dt + \widetilde{\sigma} d\mathbf{W}_t, \tag{2}$$

where $\widetilde{\kappa}$ is a diagonal matrix.

2. We look for constants $\{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}\}$ such that:

$$\begin{aligned} \widetilde{S}_{t,1} &= \alpha_{11} S_{t,1} + \alpha_{12} S_{t,2}, \\ \widetilde{S}_{t,2} &= \alpha_{21} S_{t,1} + \alpha_{22} S_{t,2}, \\ \widetilde{\kappa} &= \begin{bmatrix} \widetilde{\kappa}_1 & 0 \\ 0 & \widetilde{\kappa}_2 \end{bmatrix}. \end{aligned}$$

3. The matrix $\widetilde{\kappa}$ has eigenvalues $\widetilde{\kappa}_1$ and $\widetilde{\kappa}_2$, and the process $\widetilde{S}_{t,j}$ corresponding to the largest of these, *i.e.,* $\max\{|\widetilde{\kappa}_1|, |\widetilde{\kappa}_2|\}$, will have the strongest exposure to the mean-reverting process and hence should contain the most trading-relevant information (that is it will generate the best trading signal).

Empirics of Pairs Trading (Contd ...)

1. Illustrative Example: Estimation of the relationship between INTC and SMB during November 1, 2013.
2. The sampling is done using the mid-price and estimate the process at regular intervals (every 5 seconds).
3. Fit the discrete version of the model in (1) and use it to compute the values of the transformed model in (2) in order to build the trading signal.
4. For the discrete version of the model (1) the Vector Autoregressive Process (VAR):

$$\Delta \mathbf{S}_t = \mathbf{A} + \mathbf{B}\Delta \mathbf{S}_{t-1} + \epsilon_t,$$

where $\mathbf{S}_t = [S_{t,INTC} \;\; S_{t,SMB}]^\top$.

	A	B			
		$\Delta S_{t-1,INTC}$		$\Delta S_{t-1,SMH}$	
$\Delta S_{t,INTC}$	0.011	0.997	***	0.002	
$\Delta S_{t,SMH}$	0.035	0.003		0.998	***

Table 3.11 Estimated parameters of VAR (*** significant at 1% level).

Figure: Table 3.11

Empirics of Pairs Trading (Contd ...)

Model parameters:

$$\begin{aligned}
\kappa &= \begin{bmatrix} 0.003 & -0.002 \\ -0.003 & 0.002 \end{bmatrix}, \\
\theta &= \begin{bmatrix} 24.30691 \\ 40.91387 \end{bmatrix}, \\
\widetilde{\kappa} &= \begin{bmatrix} 0.0047 & 0 \\ 0 & 0.0007 \end{bmatrix}, \\
\widetilde{S}_t &= \begin{bmatrix} 0.682 & 0.547 \\ -0.731 & 0.837 \end{bmatrix}.
\end{aligned}$$

Figure 3.6

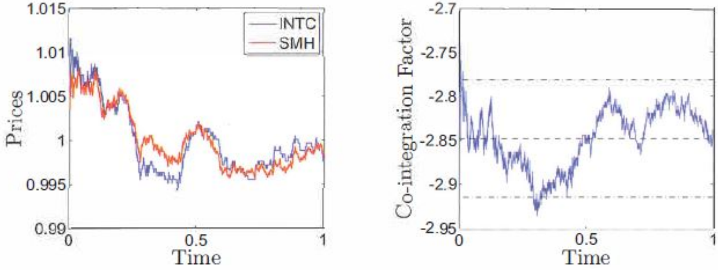

Figure 3.6 INTC and SMH on November 1, 2013: (left) midprice relative to mean midprice; (right) co-integration factor. The x-axis is time in terms of fractions of the trading day. The dashed line indicates the mean-reverting level; the dash-dotted lines indicate the 2 standard deviation bands.

Figure: Figure 3.6

Empirics of Pairs Trading (Contd ...)

1. Figure 3.6: The price process for the two assets is displayed in the left panel.
2. In the right panel the price process for $\widetilde{S}_1$ is displayed.
3. (Not shown here): Much stronger mean-reversion for $\widetilde{S}_1$ than for $\widetilde{S}_2$.