MA668: Algorithmic and High Frequency Trading
Lecture 14

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

Daily Volume and Volatility

1. Price level (and the asset's return) over the course of a whole day is difficult to predict.
2. For short time horizons:
   - A. These prices have fat tails.
   - B. Are subject to rapid changes (where the speed of change of price levels depends on the frequency with which the asset is traded).
   - C. These changes tend to cluster in time.
   - D. Are more likely than not to return to their previous level.
3. In contrast, trading activity, usually measured using volume (either in number of shares or the value of shares traded) has a different dynamic structure that has important ramifications for the way we look at market data.
4. In the words of Andersen and Bondarenko (2014):
   "`Since volume and volatility are highly correlated and display strong time series persistence, any variable correlated with volatility will, inevitably, possess non-trivial forecast power for future volatility. This is true for bid-ask spreads, the quoted intensity, the transaction count, the (normalized) trading volume ...`"

Daily Volume and Volatility (Contd ...)

1. Consideration of the empirical aspects of some of the variables associated with volatility.
2. We look at the relationship between volume and volatility using robust regression for the four main assets (ISNS, FARO, MENT, AAPL) with daily volume as the dependent variable.
3. The estimation of two models is carried out using robust OLS.
   - A. LHS variable: Involves log of the number of shares traded on each trading day of 2013 in the form $\log(1 + Q_t)$ (here $(1 + Q_t)$ is used, since, for some assets, e.g., ISNS, there are some days with no trading and zero volume).
   - B. RHS variables: Variables are the same as used in the models for intraday returns.

Daily Volume and Volatility (Contd ...)

1. Model M1:

$$\begin{aligned}\log(1+Q_{t,j}) &= \alpha + \beta_{1,j}\log(1+Q_{t-1,j}) + \beta_{2,j}SPY_t + \beta_{3,j}VIX_t \\ &+ \beta_{4,j}r_{t,j} + \beta_{5,j}OF_t + \epsilon_j.\end{aligned}$$

2. Model M2:

$$\begin{aligned}\log(1+Q_{t,j}) &= \alpha + \beta_{1,j}\log(1+Q_{t-1,j}) + \beta_{2,j}SPY_t + \beta_{3,j}VIX_t \\ &+ \beta_{4,j}r_{t,j} + \beta_{5,j}OF_t + \beta_{6,j}(SPY_t)^2 + \beta_{7,j}(VIX_t)^2 \\ &+ \beta_{8,j}HL_t + \beta_{9,j}(r_t)^2 + \epsilon_j.\end{aligned}$$

3.
   - A. SPY_t: Intraday return of SPY ETF (Market-wide volatility).
   - B. VIX_t: Intraday return on VIX (Volatility of volatility).
   - C. r_t: Intraday return on asset (Intraday volatility).
   - D. OF_t: Net order flow for the day.
   - E. HL_t: Asset's price range $[\max P_t - \min P_t]$ during the day (Day's price volatility).

Table 4.1

Variables	ISNS M1	ISNS M2	FARO M1	FARO M2	MENT M1	MENT M2	AAPL M1	AAPL M2
constant	**6.47**	**5.40**	**4.88**	**5.22**	**7.77**	**7.70**	**5.46**	**9.66**
$\log 1 + Q_{t-1}$	**0.22**	**0.22**	**0.58**	**0.52**	**0.41**	**0.39**	**0.67**	**0.38**
SPY (%)	0.04	0.19	-0.17	**−0.21**	0.03	0.07	0.00	0.03
VIX (%)	0.02	0.02	0.00	-0.01	0.01	0.02	0.00	-0.00
r_t	-0.01	-0.03	**0.05**	0.04	0.06	0.01	0.01	-0.01
Order Flow	0.02	0.04	0.00	-0.00	-0.01	-0.00	-0.00	0.00
SPY^2	—	-0.01	—	0.09	—	-0.11	—	-0.03
VIX^2	—	0.00	—	-0.00	—	0.00	—	0.00
HL-volat	—	**0.32**	—	**0.10**	—	**0.20**	—	**0.28**
r_t^2	—	-0.01	—	0.01	—	-0.02	—	-0.02
Adj R	0.03	0.18	0.30	0.50	0.17	0.24	0.38	0.65

Table 4.1 Robust OLS regression of intraday volume (Bold: 5% significance)

Figure: Table 4.1

Table 4.1 (Contd ...)

1. No evidence of significant effects from market variables (SPY or VIX) nor from order flow on volume.
2. No effect of day's intraday returns (except FARO, in case of M1, which disappears once better proxies for intraday volatility are included).
3. Observation: Substantial support for the statement due to Andersen and Bondarenko:
   - A. `Volume seems to have significant time series persistence as evidenced by the common, positive and significant coefficient on the last period's volume, in case of both M1 and M2, for all assets.`
   - B. `Positive and significant ''correlation'' with volatility as measured by HL-volatility (in M2, for all assets).`
   - C. `Volatility, as measured by the square of intraday returns, seems statistically insignificant, in the presence of HL-volatility.`

Intraday Activity: Figure 4.1

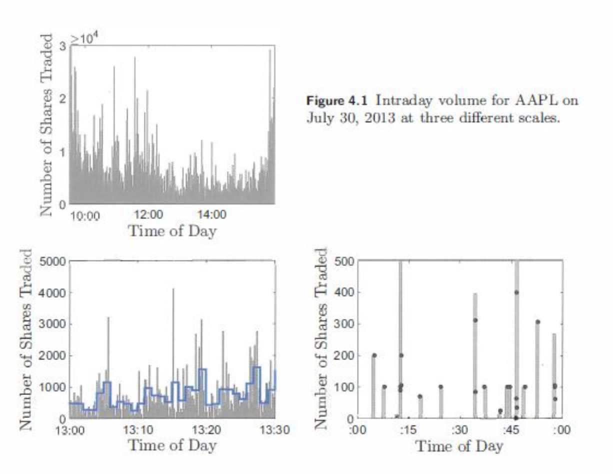

Figure 4.1 Intraday volume for AAPL on July 30, 2013 at three different scales.

Figure: Figure 4.1

Intraday Activity: Figure 4.1 (Contd ...)

1. We examine some other well known patterns of intraday volume.
2. Figure 4.1: Volume (number of shares traded in NASDAQ) for three different time scales over the course of a single trading day for AAPL.
3. Top Panel: Shows results over the whole day when volume is aggregated in one minute buckets (although the volumes for the first and the last couple of minutes are off the scale).
   - (A) Striking Characteristic 1: Peaks at the beginning and at the end of the day. Third peak is at noon, which will later be concluded as being representative of a pattern that is specific to this trading day (and not generic to the trading of the asset).
   - (B) Striking Characteristic 2: Large variability in volume. A computation of the descriptive statistics results in $\mu = 6,898$ shares, $\sigma = 7,014$ shares, Q1 $= 3,299$ shares, Median $= 5,349$ shares and Q3 $= 8,039$ shares. Further, the Skewness and Kurtosis were computed to be 6.14 and 60.83, respectively.

Intraday Activity: Figure 4.1 (Contd ...)

1. Bottom Two Pictures: Pattern of volume during a 30-minute and a 1-minute window in the middle of the day.
2. Left Panel: Compares volume aggregated in 10-seconds bucket with their 1-minute average (the thick line).
3. Substantial variation with large peaks of trading mixed in with periods of relative calm.
4. Right Panel: The pattern is even more striking. For a single minute of the day, the gray columns identify volume aggregated to one second, while the large dots represent volume aggregated at 20 ms (which is almost equivalent to plotting individual transactions).
5. Transactions seem randomly distributed over the minute, and it is not obvious whether the clustering of changes in prices discussed earlier, is also taking place at this time scale.
6. Transactions also appear to be happening in multiples of 100 shares.

Intraday Activity

1. The fact that transactions occur at round quantities is an institutional feature.
2. Market designers and regulators differentiate between "odd", "even", "mixed" and "round lots".
   - A. `Round lot is a message or transaction involving units of even lots (an even lot is 100 lots).`
   - B. `Odd lots are trades smaller than an even lot and can be more expensive to trade (in terms of fees/commissions).`
   - C. `Mixed lots are transactions which include both round lots and odd lots.`
   - D. `Odd lots are sometimes not even displayed on the consolidated tape.`

Intraday Activity (Contd ...)

1. Many algorithms are based on or linked to volume e.g., some execution algorithms may require that MOs sent to the exchange do not exceed a percentage of what other market participants are trading at that point of time.
2. Some algorithms are designed to trade in a given direction, buy or sell the asset, while targeting a given percentage of the market. These algorithms are known as Percentage of Volume (POV) algorithms.
3. Moreover, volume plays a very important role in determining execution cost benchmarks.
4. One of the most important of these is the Volume Weighted Average Price (VWAP).

Intraday Volume Patterns: Figure 4.2

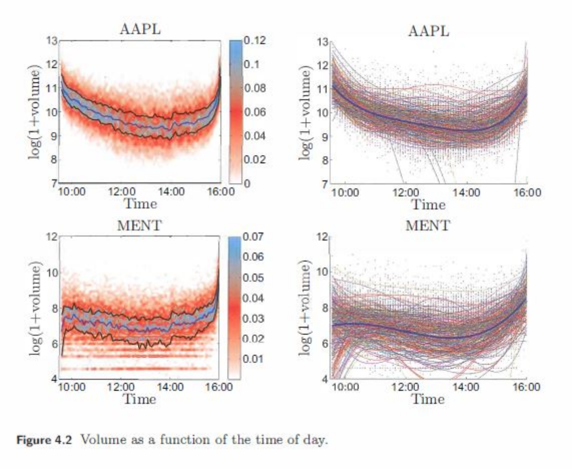

Figure: Figure 4.2

Intraday Volume Patterns: Figure 4.2 (Contd ...)

1. Figure 4.2: Shows heat-maps (left panes) of daily volume for the year 2013 for AAPL and MENT.
2. The heat-maps are generated by first bucketing traded volume into five-minute windows throughout the day, for every day of the year.
3. Then for each five minute bucket, we compute the distribution of volume.
4. The heat-map is a visualization of the collection of these distributions for each five minute bucket all at once.
5. In the figure, the colored lines represent the first and the third quartiles, as well as the median.

Intraday Volume Patterns: Figure 4.2 (Contd ...)

1. We observe that the volume for AAPL, is very large at the beginning of the day, and it gradually slows down until around 14:00.
2. At 14:00 there is small surge in activity which slowly builds up and accelerates during the last half hour of the trading day, peaking at the close.
3. Reasonable hypothesis for the 14:00 surge is that at that time there are more announcements [a] than is the norm, and these announcements tend to generate greater volume.
4. Figure for July 30: Apart from the usual peaks at the beginning and at the end of the day, there is also a peak of trading volume around noon, which is (as already noted) atypical for this asset.
5. To explain the peaks at the beginning and at the end of the day, we need to hypothesize about the factors that drive volume.

[a] For example, the monthly Treasury Budget is announced at that time of the day.

Intraday Volume Patterns: Figure 4.2 (Contd ...)

1. A common hypothesis, which is made here also, is that new information generates greater volume.
2. In addition to the 14:00 surge, this would also explain why there is such a large volume "at" and "just after" the opening, since overnight news is gradually into the prices during that time.
3. However, this hypothesis does not explain the magnitude of increase in volume at the end of the day, as there does not exist an unusually large number of announcements at that time.
4. Another possible explanation for this peak in trading at the end of the day: Traders who have not been able to meet their liquidation targets will accelerate trade execution as the market approaches its time to close.
5. A second, not unrelated possible explanation is that traders may prefer to postpone non-urgent executions towards the end of the day, when execution costs are lower (details later).

Intraday Volume Patterns: Figure 4.2 (Contd ...)

1. Right panel of Figure 4.2: Presents the functional data analysis (FDA) approach to viewing the data.
2. For every trading day, we regress the realized five minute volume against Legendre polynomial and plot the resulting curve as a thin line in the Figure.
3. Consequently: A smooth volume curve is generated for each day of the year.
4. Next: We plot the mean of the curves (the solid blue line) which represents the explicit (or average) trading volume throughout the day, for the compounding ticker.

Intraday Volume Patterns: Figure 4.2 (Contd ...)

1. We observe four large outliers for AAPL: Four curves that disappear off the bottom of the scale.
   - A. These represent four special days for 2013.
   - B. Three of those four were predictable, while the other was not.
   - C. Three predictable days were: July 3 (Independence Day Eve), November 29 (Thanksgiving Eve) and December 24 (Christmas Day Eve) when NASDAQ closed early (13:00) for the holidays.
   - D. Fourth outlier corresponds to August 22, 2013, when NASDAQ suffered major problems leading to about 3-hour shutdown of the market.
2. These days were excluded from the calculation for MENT.

Intraday Volume Patterns: Figure 4.2 (Contd ...)

1. In addition to identifying regularities in intraday trading patterns, the FDA curves have helped us identify outliers in the data.
2. Outliers are very important when using historical data for analysis, back-testing and designing algorithms.
3. Quite often, an outlier will have a disproportionate effect on an algorithm's probability, whether when back-testing it against historical events or running it live in the market.
4. Accordingly: It is crucial to keep track of those outliers and account for them in the design and evaluation of algorithms.
5. In the historical analysis of the intraday pattern of volume, AAPL's mean daily volume function is lower than it would have been if we had not included those four particular dates in the estimation.

Intraday Volume Patterns: Figure 4.2 (Contd ...)

1. Now we consider contrasting the intraday behaviour of AAPL with a less frequently traded asset, namely, MENT.
2. Observation: Daily peaks at the beginning and at the end of the day are also seen, but sightly modified, and distorted by the discreteness of the trading lots.
3. For MENT, the initial burst of trading is not as frequent as the AAPL.
4. In fact, there are a substantial number of days for which trading starts unusually slowly.
5. Slow days balance out the bursts of trading from other days, so that on an average the volume in early trading does not seem to differ substantially from the rest of the day.
6. Substantial amount of trading activity in MENT at the end of the day.
7. Analysis of ISNS and FARO are excluded since they are even less frequently traded.

Intrasecond Volume Pattern: Figure 4.3

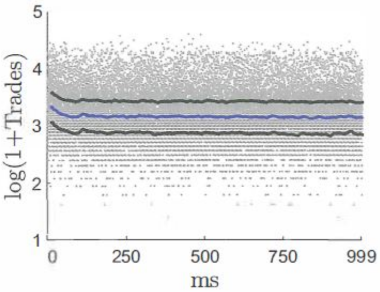

Figure 4.3 Trading pattern within a second.

Figure: Figure 4.3

Intrasecond Volume Pattern: Figure 4.3 (Contd ...)

1. Question: For time intervals finer than a second, do we observe time patterns like the ones we observed over the duration of the day.
2. For the answer, we focus only on AAPL and look at the millisecond trading pattern.
3. Figure 4.3: Displays the average number of transactions at each millisecond for each day in 2013 (AAPL), as well as Q1, Q3 and mean.
4. Observation: Hardly a persistent pattern at the millisecond level, although an initial spike is observed in the 000-020 ms range, followed by a subtle valley that ends around 100 ms point.

Intrasecond Volume Pattern: Figure 4.4

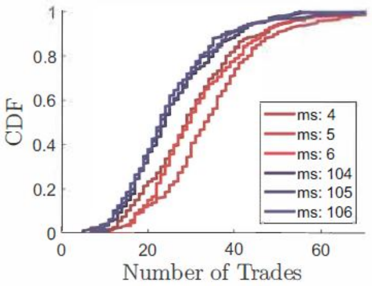

Figure 4.4 Cumulative distribution function of transaction count for specific milliseconds.

Figure: Figure 4.4

Intrasecond Volume Pattern: Figure 4.4 (Contd ...)

1. Figure 4.4: Empirical cumulative distribution function of the number of transactions ending at six different milliseconds:
   - (A) Three early ones (at 4,5,6 ms).
   - (B) Three later ones (at 104,105,106 ms)
2. Observation: Early milliseconds stochastically dominate the other three.
3. Pattern suggest that there is an unusual number of transactions that are recorded just after the exact beginning of a second.
4. Plausible explanation: There may be an unusual number of transactions that are entered (automatically) at the exact end/beginning and what we observe is the latency or clock-asynchronicity of these order (machines).