Development and Usage of Short Term Signals in Order Execution

Michael G Sotiropoulos
Algorithmic Trading Quantitative Research
Bank of America Merrill Lynch

Cornell Financial Engineering Seminar
New York, 10-Oct-2012
Signals: Concepts and Terminology

Agents: investors (institutional, retail), market makers and brokers
Process: continuous quoting and trading inside two-sided, electronic limit order books
Outcome: price discovery

Trading always involves costs.

- Explicit costs are commissions, bid-ask spread, foreign exchange fees, etc.
- Implicit costs are price impact, adverse selection and opportunity cost.

- Aggressive (market) orders pay up-front the spread cost in exchange for controlling the execution time.
- Aggressive orders deplete the order book and generate price impact.
- Passive (limit) orders reveal information and are subject to adverse selection.
- Passive orders gain up-front the spread, at the risk of been left unfilled.

Algorithmic trading research focuses on measuring and modeling costs, as well as optimally controlling the discretionary variables of trading.
Signals: Definitions and Types (I)

A trading signal is a combination of two components:

1. An **indicator function** $I(t, O_t, \ldots O_{t-h}; \theta)$.
   It depends on current and lagged market observables $O_t \ldots O_{t-h}$, and on model specific parameters $\theta$.
   The purpose of the indicator is to compute some dynamic aspect of the market, and return a short term forecast.

2. A **response function** $R(I_t, X_t; \phi)$.
   It depends on the indicator $I_t$, the current state of the order $X_t$ and strategy parameters $\phi$.
   The purpose of the response function is to generate a trading action.

Possible actions are:

- Increase/decrease the quantity of a limit order
- Update the price of a limit order
- Cross the spread
- Cancel a limit order and wait out until further updates
- Reallocate the posted quantities among trading venues (exchanges, dark pools)

Indicator + Response $\rightarrow$ Adaptive Algorithm
**Signals: Definitions and Types (II)**

Common Indicator Types:

1. **Trade autocorrelation**: trade signs are correlated within a short time scale → predictable trade direction

2. **Order imbalance**: the LOB may be too heavy on one side → predictable mid-price movement

3. **Momentum/reversion**: the price path exhibits strong trend → bet on the trend persisting or reverting

4. **Relative value**: the traded asset is cointegrated with a sector index or another asset → predictable spread movement

5. **News/Events**: the market reaction to unexpected news has a stable pattern → predictable post-event volume/volatility/alpha

6. **Volume clustering**: recent spike in trading volume is expected to create more spikes over a short horizon → predictable increase in next bucket volume

7. **Venue liquidity**: higher probability to get filled on a specific venue due to hidden volume, popularity for certain assets, fee structure, etc. → optimal routing
Signals: Real Value

Note:

- Trading signals are short lived and opportunistic (i.e. unstable in time).
- Technical indicators are the longer horizon relatives of the trading signals considered here.

Do trading signals add real value?

- In a perfectly liquid and efficient market: no.
  Prices are martingales and the best forecast for the future is the present state.
  Signals are simply white noise.

- In theoretical models for illiquid markets: maybe.
  Most signals are either due to market microstructure noise, or they are priced in the order flow.

- In real markets: yes.
  Provided that the signal is correctly identified, properly calibrated and periodically reviewed for validity.
Order Flow Models and Signals: Model Types (I)

Quantitative models for high frequency trading summarize a subset of market dynamics in a mathematical framework.

Models can be classified as:

1. **Microscopic**: sequential trading, strategic trading
   - model agent interactions (informed/noise trader, market maker)
   - optimize individual agent objectives
   - derive market clearing prices

   Examples: Roll ’84, Glosten-Milgrom ’85, Kyle ’85, MRR ’97, and more ...

2. **Macroscopic**: impact function, decay kernels
   - average over agent behavior (effective theories)
   - parametrize aggregate cost effects with simple functional forms
   - maintain the constraint of efficient markets

   Examples: zero intelligence (Smith-Farmer ’02), power law impact (Bouchaud et al. ’08)

- Macroscopic models are used to generate the target schedule of a large order.
- Trading signals are used to **opportunistically deviate** from the target schedule.
- Microscopic models allow us to distinguish between microstructure noise and genuine information.
A simple microstructure model with autocorrelated order flow and minimal strategic trading is the MRR model (Madhavan, Richardson, Roomans 1997).

Assumptions:

1. All market orders have the same quantity.

2. The trade signs follow an AR(1) Markov process, i.e.

\[
E(\epsilon_i|\epsilon_{i-1}, \epsilon_{i-2}, \ldots) = \rho \epsilon_{i-1}.
\]  (1)

The lag-1 autocorrelation of trade signs decays exponentially as

\[
C_l := E(\epsilon_i \epsilon_{i+l}) = \rho^l.
\]  (2)

3. The fundamental price is affected by the external shock \(\xi_i\) (news) and by the trade sign surprise \(\epsilon_i - \rho \epsilon_{i-1}\) as

\[
p_{i+1} = p_i + \xi_i + \theta (\epsilon_i - \rho \epsilon_{i-1}),
\]  (3)

with \(\theta\) the coefficient of price impact.

Trading mechanics: the prevailing bid and ask prices \(b_i, a_i\) are valid in the interval \([t_{i-1}, t_i)\). They get updated immediately after the arrival of trade \(i\).
Order Flow Models and Signals: Example: MRR (II)

What is a market maker to do before trade $i$ happens?
Set bid-ask prices so that there is no ex-post regret

$$b_i = p_i + \theta ( -1 - \rho \epsilon_{i-1} ) - c, \quad a_i = p_i + \theta (1 - \rho \epsilon_{i-1}) + c. \quad (4)$$

- The spread has a price impact component and a fixed/inventory cost component

$$s = a_i - b_i = 2(\theta + c). \quad (5)$$

- The mid price is the fundamental price corrected by the expected impact

$$m_i = (a_i + b_i)/2 = p_i - \theta \rho \epsilon_{i-1}. \quad (6)$$

- After trade $i$ happens the mid price moves to

$$m_{i+1} = m_i + p_{i+1} - p_i - \theta \rho (\epsilon_i - \epsilon_{i-1}) = m_i + \xi_i + \theta (1 - \rho) \epsilon_i. \quad (7)$$

- After $l$ trades have taken place

$$m_{i+l} = m_i + \sum_{j=i}^{i+l-1} \xi_j + \theta (1 - \rho) \sum_{j=i}^{i+l-1} \epsilon_j. \quad (8)$$
Order Flow Models and Signals: Example: MRR (III)

Define the impact function at lag $l$ as (Wyart, et al. 2008)

$$\mathcal{R}_l := \mathbb{E} (\epsilon_i (m_{l+i} - m_i)). \quad (9)$$

- $\mathcal{R}_l$ measures the mid-price impact of trade $i$ over a horizon of $l$ time steps.
- It is easily computed from eq. (8) as

$$\mathcal{R}_l = \theta \left(1 - \rho^l\right). \quad (10)$$

- For $\rho > 0$ (the case in practice), the impact increases from $\mathcal{R}_1 = \theta (1 - \rho)$ to $\mathcal{R}_\infty = \theta$.

Conclusions:

1. Positive correlation among trade signs leads to increased long term impact

$$\mathcal{R}_\infty = \frac{1}{1 - C_1} \mathcal{R}_1. \quad (11)$$

2. The spread is a linear function of the long term impact

$$s = 2\theta + 2c = 2\mathcal{R}_\infty + 2c. \quad (12)$$

3. In the absence of price drift (alpha) the long term impact cost ($\mathcal{R}_\infty$) of market and limit orders is the same (order duality).
   The long term total cost of market and limit orders differs only by the fixed spread cost $2c$. 


Order Flow Models and Signals: Lessons

Models provide a healthy criticism about signals because:

- Microstructure models generate transaction prices that are not martingales.
- Deviation from martingale behavior does not necessarily lead to a meaningful signal. For example:
  - Negative autocorrelation of trade prices in the Roll ’84 model is due to bid-ask bounce
  - Positive autocorrelation of trade signs in the MRR ’97 model cannot be exploited, it is priced in the order flow

To develop meaningful signals we need to:

1. Check the underlying dynamics that motivate the indicator.
2. Calibrate the indicator parameters and time window.
3. Tune the response function via back-testing or randomized testing.
Having chosen a market observable $O_t$, how far back do we look to construct the indicator $I_t$?

Approximating the LOB as a simple queue with constant service rate, we define the **queue time** as the average time it takes to move the mid-quote price by depleting the bid or the ask side

$$\tau_q := \frac{\text{AvgBidSize} + \text{AvgAskSize}}{\text{AvgTrdSize}}.$$  \hspace{1cm} (13)

- Long queue stocks have thick LOBs relative to the typical trade size.
- The spread (tick size) of long queue stocks is large relative to the stock price, so limit orders pile up at the top of the book.
- Short queue stocks are typically liquid, with the top of the book updating quickly.

**NOTE:** the units of $\tau_q$ is number of trades (tick time, not clock time).
The time window used to construct an indicator may be defined as:

1. The queue time translated into wall clock units (with some zoom factor $z$)

$$\tau_w = \max \left( \tau_f, \min \left( \tau_c, z \frac{T_{day}}{N_{trd}} \tau_q \right) \right) ,$$

(14)

2. The time interval that contains on average $n$ number of trades

$$\tilde{\tau}_w = \max \left( \tau_f, \min \left( \tau_c, \frac{T_{day}}{N_{trd}} n \right) \right) ,$$

(15)

where

$\tau_f, \tau_c$ user-specified floor and cap

$T_{day}$ time length of the continuous trading session

$N_{trd}$ average number of trades per day

The quantity $N_{trd} / T_{day}$ is the market average speed of trading.
Signal Examples: Time Scales and Weights (III)

Distribution of WndQSize (window size with $\tau_q$ trades on average) and WndSize (window size with 16 trades on average) for S&P 500 stocks.

- Each time is an average across all trading days in October 2011
- Configuration: $\tau_f = 1$ sec, $\tau_c = 3$ mins

- WndQSize outliers: TLAB (4.5 mins), WPO (3.5 mins)
- WndSize outliers: WPO (23 mins), GAS (3.4 mins)
Signal Examples: Time Scales and Weights (IV)

Same distributions for Russell 3000 stocks.

- Each bar is 60 secs wide
- The cap is 3 mins

- Proportion of stocks that hit the WndQSize cap: $44/2941 = 1.5\%$
- Proportion of stocks that hit the WndSize cap: $215/2941 = 7.3\%$
Signal Examples: Time Scales and Weights (V)

Within $\tau_w$ we use exponential moving averages (EMA) for the lagged observations.

The standard definition of the EMA $M_t$ of a quantity $O_t$ is

$$M_t = aO_t + (1 - a) M_{t-1}$$

(16)

The coefficient $a$, between 0 and 1, is called the “smoothing factor” (bad name).

- $a \rightarrow 1$ the EMA discounts the past and tracks the present more closely.
- $a \rightarrow 0$ the EMA discounts the present and it is smoother.
- The EMA is 86% determined by the last $(2/a)$ observations.

Define an exponentially weighted time distance between trades at $t_{i-1}$ and $t_i$

$$w_i = e^{-(t_i-t_{i-1})/\tau_w}$$

(17)

Define the smoothing factor at trade time $t_i$ as

$$a_i = 1 - w_i = 1 - e^{-\Delta t_i/\tau_w}.$$  

(18)

Assuming that trades arrive at a constant speed $N_{trd}/T_{day}$ then $\Delta t_i \approx T_{day}/N_{trd}$ and from eq. (14),

$$a \approx 1 - e^{-1/\tau_q}.$$  

(19)

For $\tau_q \gg 1$ the smoothing factor becomes

$$a \approx \frac{1}{\tau_q},$$

(20)

i.e. the indicator is determined by the last $2\tau_q$ trades.
Signal Examples: Trade Sign Autocorrelation (I)

- This signal exploits the persistence of the order flow.

Why is it expected to work?

- Use the Lee-Ready algorithm to sign the trades on the tape (1: BUY, -1: SELL)
- Compute the autocorrelation function (ACF), i.e. the correlation between
  - every trade sign and the next trade sign (lag = 1)
  - every trade sign and the sign two trades after (lag = 2)
  - every trade sign and the sign $h$ trades after (lag = $h$)

If trade signs arrive independently the ACF for all lags (except lag 0) should have a mean of zero.
Signal Examples: Trade Sign Autocorrelation (II)

Log-log plot of the autocorrelation of trade signs (using the Lee-Ready algorithm for signing trades). Note the strong autocorrelation for a significant number of lags.

- Primary reason is order splitting (Tóth, et al. 2011).
- Power law decay $\rho \propto h^{-\gamma}$, with $\gamma = 0.50$ (MSFT) and $\gamma = 0.65$ (BEAM).
Signal Examples: Trade Sign Autocorrelation (III)

Calculation of the signal (Almgren 2006)

1. For each trade define its “askness” \( a \) and “bidness” \( b \) as the distance of the transaction price from the bid (res. ask) in units of spread

\[
\begin{align*}
    a &= \min \left( \left( \frac{P - P_b}{P_a - P_b} \right)^+, 1 \right) ; \quad b = \min \left( \left( \frac{P_a - P}{P_a - P_b} \right)^+, 1 \right). 
\end{align*}
\]

(21)

By construction, \( a + b = 1 \). A trade that hits the ask side (BUY) has \( a = 1, \ b = 0 \).

2. At each trade time \( t_n \) compute the moving average of askness and bidness over a window of size \( \tilde{\tau}_w \) as

\[
\begin{align*}
    A_n &= \frac{1}{\tilde{\tau}_w} a_n + w_n A_{n-1} ; \quad B_n = \frac{1}{\tilde{\tau}_w} b_n + w_n B_{n-1},
\end{align*}
\]

(22)

with exponentially decaying weights \( w_n = e^{- (t_n - t_{n-1}) / \tilde{\tau}_w} \).

3. Normalize the moving averages by half the average trading speed

\[
\begin{align*}
    \bar{A}_n &= \frac{2A_n}{N_{trd} / T_{day}} , \quad \bar{B}_n = \frac{2B_n}{N_{trd} / T_{day}}
\end{align*}
\]

(23)

An algorithm that tries to minimize impact cost will use the signal as follows:

- For a BUY (SELL) order trade faster when \( \bar{B}_n (\bar{A}_n) \) is higher.
Signal Examples: Trade Sign Autocorrelation (IV)

What does the signal mean?

- If the order flow was balanced within the window $\tau_w$, then $\bar{A}_n \approx \bar{B}_n \approx 1$. Half of the trades on average should be BUY and half SELL.
- If we are posted on the bid side and $\bar{B}$ is high, there is a lot of SELL market orders, so we should increase our participation rate (response function in the trading system).

This is one of the signals used by the BAML Instinct® algorithm.
Signal Examples: Trade Sign Autocorrelation (V)

A real order:

<table>
<thead>
<tr>
<th>Algo</th>
<th>BAML Instinct®</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>STT</td>
</tr>
<tr>
<td>Side</td>
<td>SELL</td>
</tr>
<tr>
<td>TargetPct</td>
<td>20%</td>
</tr>
<tr>
<td>IvIReturn</td>
<td>15.4 bps</td>
</tr>
<tr>
<td>Slippage</td>
<td>-6.2 bps</td>
</tr>
</tbody>
</table>

**Px**

<table>
<thead>
<tr>
<th>13:51</th>
<th>13:53</th>
<th>13:56</th>
<th>13:58</th>
<th>14:00</th>
<th>14:02</th>
</tr>
</thead>
<tbody>
<tr>
<td>38.80</td>
<td>38.86</td>
<td>38.92</td>
<td>38.99</td>
<td>39.05</td>
<td>39.05</td>
</tr>
</tbody>
</table>

**Vlm (1000s)**

<table>
<thead>
<tr>
<th>13:51</th>
<th>13:53</th>
<th>13:56</th>
<th>13:58</th>
<th>14:00</th>
<th>14:02</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>3.8</td>
<td>7.2</td>
<td>10.6</td>
<td>14.1</td>
<td></td>
</tr>
</tbody>
</table>

**Total Vlm Instinct Signal**

<table>
<thead>
<tr>
<th>13:51</th>
<th>13:53</th>
<th>13:56</th>
<th>13:58</th>
<th>14:00</th>
<th>14:02</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.81</td>
<td>1.63</td>
<td>2.44</td>
<td>3.25</td>
<td></td>
</tr>
</tbody>
</table>

**Target %**

<table>
<thead>
<tr>
<th>13:51</th>
<th>13:53</th>
<th>13:56</th>
<th>13:58</th>
<th>14:00</th>
<th>14:02</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>7.3</td>
<td>14.7</td>
<td>22.0</td>
<td>29.4</td>
<td></td>
</tr>
</tbody>
</table>

**Realized %**

<table>
<thead>
<tr>
<th>13:51</th>
<th>13:53</th>
<th>13:56</th>
<th>13:58</th>
<th>14:00</th>
<th>14:02</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Signal Examples: Order Imbalance (I)

Replace the mid-quote by a price that takes into account the imbalance between bid and ask sizes

**Microprice:** a LOB imbalance signal defined as a “center of mass” price within the spread

\[
P_{\text{micro}} := P_b \frac{Q_a}{Q_b + Q_a} + P_a \frac{Q_b}{Q_b + Q_a}.
\]  

(24)

Response function:

- Buy order: cross the spread when \( P_{\text{micro}} > P_a - k (P_a - P_b) \).
- Sell order: cross the spread when \( P_{\text{micro}} < P_b + k (P_a - P_b) \).

Tune \( k \) based on empirical studies of order performance.
Signal Examples: Order Imbalance (II)

Joint evolution of the NBBO and the MicroPrice for MSFT and BEAM.

- Notice how MicroPrice anticipates the shift of the NBBO level for MSFT (less clear for BEAM).
Signal Examples: Order Imbalance (III)

MicroPrice crossing is a horizontal feature of the BAML limit order model. It triggers opportunistic spread crossing of the child order.

Note:

- MicroPrice-like state variables can determine the probability of up-down move of the next price innovation in Markov models for the LOB (Cont, de Larrard 2012).
- Spread crossing is costly. The cut-off must be carefully calibrated.
- Latency effects may significantly reduce the benefit of an order imbalance indicator (Stoikov, Waeber 2012).
Signal Examples: Intraday News (I)

Facts:
- Corporate and macro news affect intraday trading volume and volatility.
- Regularly scheduled releases (earnings, FOMC meetings) are considered “special days” and trading systems load special day statistics.
- The problem is to assess and react to unscheduled news intraday.

Methodology:
- Use linguistic analysis to interpret and score news items, i.e. map them to numerical indicators
- Define a meaningful and robust scoring system
- Create a real-time feed that provides the stream of scores to the trading engines

 Providers:
- News feeds scoring and distribution is a fast maturing industry
- Main providers: Thomson Reuters, Bloomberg, Dow Jones
Signal Examples: Intraday News (II)
Example from Thomson Reuters.

Some of the indicators provided by the feed are

1. Item type (article, alert, append)
2. Relevance (between 0 and 1)
3. Sentiment (±1)
4. Positive/Neutral/Negative weight (the three weights sum up to 1)
Signal Examples: Intraday News (III)

To assess the impact of news on trading volume we compute a given day’s news-conditional volume as (Gross-Klussmann, Hautsch 2009)

\[ TV'_{m,k} = \frac{\sum_{i \in k} V_i P_i}{\sum_{i \in k} P_i} \]  

and its historical \( D \)-day average as

\[ \hat{T}V_{m,k} = \left( \sum_{j=d-D}^{d-1} TV_{m,k,j} \right) / D \]  

where

- \( m \) news item
- \( k \) the \( k \)-th time interval of fixed size \( \Delta T \) after the arrival of news item \( m \)
- \( V_i, P_i \) the volume and price of trade \( i \) within interval \( k \)
- \( d \) the day index

Finally, the normalized news-conditional volume is computed as

\[ TV_{m,k} = \frac{TV'_{m,k}}{\hat{T}V_{m,k}} \]
**Signal Examples: Intraday News (IV)**

Below we plot the normalized conditional volume averaged over all news items as a function of the time interval. The sample contains:

- Time period: Jan-Jun 2011
- Stock universe: FTSE 100 and FTSE ALL SHARE
- News type: ALERT or ARTICLE
- News relevance: REL = 1
- Sentiment: POS > 0.85 or NEG > 0.85

FTSE 100 on the left and FTSE ALL SHARE on the right
Signal Examples: Intraday News (V)

Intraday volume responds to news arrival as follows:

- It peaks at 1.5 the historical value around news arrival
- The peak has a $t$-test score of 2 (it is significant)
- The peak is not a jump, it has a 10 min width on either side, with a longer right tail (insiders?)
- The total daily and closing auction volumes seem unaffected by intraday news; news create a redistribution effect

Practical benefits:

- Integrate the news feed indicators with the real-time volume prediction model
- Calibrate the response to news relative to the default prediction model.
- Repeat exercise for volatility

So far we have found no significant link between news arrival and price movement (short-term alpha).
Summary: Usage in Trading Systems

Coarse-grained view of an algorithmic trading stack

- Order
  - AlgoChooser
    - Historical Statistics
    - Impact Model
  - Scheduler
    - Historical and Real Time Statistics
    - Impact Model
  - Slicer
    - Trading Signals (Imbalance, Autocorrelation, ...)
  - Router
    - Order Book Statistics
- Markets
Summary: Conclusions

- Signals are used extensively by execution service providers and high frequency firms, although the objectives and risk preferences may differ.
- Implementation and validation of short-term signals is a non-trivial task.
- Backtesting is valuable, but continuous monitoring and tuning is even more important.
- Weighted mixtures of signals, or “super-signals”, can be useful in markets with frequent regime changes.
References


All statements in this presentation are the author’s personal views and not necessarily those of Bank of America Merrill Lynch.