Are HFTs anticipating the order flow? Cross-venue evidence from the UK market
FCA Occasional Paper 16

Matteo Aquilina and Carla Ysusı

Algorithmic Trading: Perspectives from Mathematical Modelling Workshop
London, 1st March 2017
Disclaimer

Occasional Papers contribute to the work of the FCA by providing rigorous research results and stimulating debate. While they may not necessarily represent the position of the FCA, they are one source of evidence that the FCA may use while discharging its functions and to inform its views. The FCA endeavours to ensure that research outputs are correct, through checks including independent referee reports, but the nature of such research and choice of research methods is a matter for the authors using their expert judgement. To the extent that Occasional Papers contain any errors or omissions, they should be attributed to the individual authors, rather than to the FCA.
(Main?) criticism of HFTs

- They only intermediate trades that would take place in any case;
- They earn (near) risk-free profits and are – in fact – a ‘tax’ on other participants;
- Popularised in *Flash Boys* in 2014 but existed beforehand.
The mechanism...

1) Non-HFT sends a marketable order to a venue.
2) HFT sees the execution and can infer that a similar order is going to arrive at a different venue “soon”.
3) HFT then exploits its speed advantage and trades in front of the slower non-HFT in other venues (earning near riskless rents).
What do we mean by ‘soon’?

We investigate two different timeframes:

1) Do HFTs exploit small (millisecond) speed advantages at the expense of other market participants?

2) Can HFTs anticipate the order flow over longer timeframes (tens of seconds)?
Our Dataset

Order book data:
– 2 trading venues and 5 order books (LSE, Bats/ChiX). Includes both ‘lit’ and ‘dark’ books in BATS/ChiX. They cover around 85% of FTSE trading (only significant missing venue is Turquoise).

– 120 stocks (60 FTSE100, 60 FTSE250) for the whole of 2013.

– All messages recorded by venues’ matching engines (entered, amended, deleted and traded orders).

– Information on: date, time, bid/ask, price, quantity (disclosed and undisclosed), counterparties, basic order type (market/limit).

– We use data from 8:00am to 4:30pm and exclude auctions and dark books.
Different types of participants

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFTs</td>
<td>26 firms</td>
</tr>
<tr>
<td>“Pure” Non-HFTs</td>
<td>138 firms</td>
</tr>
<tr>
<td>“Other” (i.e. Investment Banks)</td>
<td>14 firms</td>
</tr>
</tbody>
</table>

HFT: Based on firm’s name. A list from a previous project was updated with help of supervision colleagues. Other: likely to be co-located and/or have DMA/SA.
How much and where ‘pure’ non HFT trade….

<table>
<thead>
<tr>
<th></th>
<th>HFTs</th>
<th>Other</th>
<th>“Pure” Non-HFTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of firms</td>
<td>14% of firms 30% of £volume traded</td>
<td>8% of firms 64% of £volume traded</td>
<td>78% of firms 6% of £volume traded</td>
</tr>
</tbody>
</table>

37 firms
5% of total trading volume (£)
80% of non-HFTs trading volume (£)
Analysis of near simultaneous orders
Methodology

Near simultaneous = marketable, sent by the same participant, for the same instrument, with the same direction, arrive at different venues within ‘a few’ milliseconds.

We analyse both “pairs” of orders and “bursts” of orders.

HFTs have to either trade in the correct direction or cancel/amend their orders in the ‘correct’ way.

Clock synchronisation issues complicate matters further (reliable within a venue but need a 1 millisecond tolerance across venues).

Our benchmark: any same direction marketable order for the same instrument (i.e. not necessarily those of the same non-HFT).
A bit more on clock synchronisation issues...

Within a single venue we know what event took place ‘before’ and ‘after’...

...not so when there are two or more venues involved as the clocks are not perfectly synchronised and there is latency (this is true for events that are timestamped with up to 1 millisecond difference)

When analysing “pairs” of orders we assume that anticipation took place only if we are 100% sure that the HFTs acted ‘before’ (which in practice implies “looking back in time”)

Financial Conduct Authority
Aggressive trading within a day: early and late trading most common.

Median frequencies per second
When are participants sending aggressive orders? HFTs are quicker than non-HFTs. HFTs seem to lead non-HFTs aggressive orders. Interestingly, HFTs seem to trade aggressively again after non-HFTs.

All TVs: Periodicity of Aggressive Orders within a Second
HFT activity between pairs of near simultaneous orders: activity is low overall and never (statistically) significantly larger than the benchmark.
HFT activity between **strings** of near simultaneous orders: activity is low overall and never (statistically) significantly larger than the benchmark.

*In TVs and securities where "pure" non-HFTs send simultaneous marketable orders*
Analysis of longer timeframes
Methodologies

Brogaard (2010):

Compare the frequencies of trades initiated by HFTs that precede different size non-HFTs initiated trades.

Hirschey (2013):

Compare cumulative net marketable imbalances and returns of HFTs and non-HFTs.
Methodology of Brogaard (2010)

- If HFT are detecting large orders we would expect them to increase their activity just before a large trade.

- To detect if this occurs on a systematic basis:
  - For each stock, create bins based on the trade size for non-HFT initiated trades, i.e. classify small/large non-HFT initiated trades.
  - Compare the average percentage of HFT initiated trades prior to the non-HFT initiated trade. If HFT can detect the larger trades, we would expect HFTs will be more active before larger trades.
  - We improved the methodology by considering if an HFT was the counterparty of the non-HFT initiated trade.

How big is the trade?

Are HFTs increasing their activity just before a large trade?

The same HFT that traded against the non-HFT initiated trade

HFT is the counterparty

Non-HFT initiated trade

Any HFT
Results: checking for the identity of the HFT the non HFT traded with we see that the HFT was more active the larger the size of the trade.
Methodology of Hirschey

If HFTs were anticipating non-HFTs order flow, we would expect aggressive HFT buying (selling) preceding aggressive non-HFT buying (selling).

1. Each second, identify the main stocks being aggressively bought (sold) by HFTs (90th percentile) by calculating HFTs Net Marketable Imbalances.
   - Net Marketable Buying Imbalance = #shares in buyer-initiated trades - #shares in seller-initiated trades

2. Every second, create a portfolio with those stocks that HFTs buy (sell) most aggressively.

3. Compare HFTs vs non-HFTs Net Marketable Imbalances for this portfolio around the time of the buying(selling) pressure.

4. Calculate returns before/after the second of buying/selling pressure.
   - Are prices moving in HFTs’ favour?

* Traded volumes are normalised to be able to compare net marketable imbalances between stocks.
HFTs’ net marketable buying(selling) leads non-HFTs’ net marketable buying(selling).
Restricted sample: only for those securities traded by pure non HFTs and at times in which they are active.
HFTs make profits: stocks they bought aggressively have positive returns, while those they sold aggressively have negative returns.

Cumulative midprice returns around periods when HFTs buy/sell aggressively.
Conclusions

1) No evidence for the hypothesis that HFTs are able to anticipate near-simultaneous orders routed to different venues.

2) Patterns consistent with HFTs anticipating the order flow over longer time periods (but could be faster reaction to news)

3) Welfare implications of order anticipation are unclear: they depend on what type of orders are being anticipated.