A proposed research methodology for High Frequency Trading Cost-Benefit Analysis

BY ASHRAF NOUMIR & FREDERI VIENS
Introduction

- HFT has no clear definition. SEC concept release (2010) identified 5 characteristics of HFT
  - Use of high speed sophisticated programs
  - Use of co-location services
  - Extremely short time horizon for trades
  - High order to trade ratio
  - No inventory held overnight.


- With High Frequency Trading (HFT) currently exceeding 50% of the trading volume of U.S. Listed stocks, it became a significant component of current market structure (SEC, 2014).

- The flash crash of May 6, 2010 raised public and academic concerns about whether HFT has positive or detrimental effect on market quality.

- Policy makers became more interested in proposing policies directed toward HFT.
Motivation

- Theoretical and empirical disagreement with respect to the role of HFT in market quality.
  - Theoretically (Zhang, 2010)
    - Proponents argue that
      - HFT lower stock price volatility through its market making activity
      - HFT reduce transaction cost through liquidity making activity
      - Increasing liquidity through HFT enable institutional investors to base their trading on the fundamentals of companies’ performances.
    - Opponents argue that
      - HFT increase volatility through the interaction between HFT and institutional investors (flash liquidity, front running, creating price momentum)
Empirically:

- Brogaard (2012) found that HFT increase liquidity and lowers volatility.
- In reply to concerns whether HFT increase execution cost (mainly for institutional investors), Brogaard et al. (2014) empirically found that there is no significant association between HFT and execution cost.

On the other hand,

- Zhang (2010) found positive correlation between HFT and stock price volatility with stronger correlation in periods of high market uncertainty.
- In addition, he showed that HFT decrease price discovery which he attributed to the notion that HFT hinders market’s ability to incorporate fundamentals into asset prices.
- Bershvova and Rakhlin (2013) empirically found that HFT is positively correlated with volatility.
We believe that these mixed findings might be partly attributed to the frequency (time frame) of data studied.

This study can be seen as the cross road between Brogaard (2012) [high frequency analysis] and Zhang (2010) [low frequency analysis]

Our objective is thus to show empirically evidence of the short and long term effect of HFT on the market, specifically to get a more detailed picture of the dynamics of HFT and their effects on market quality.

A lack of empirical studies to determine the presence of any relationship between the emergence of HFT and U.S. financial sector human resources

Analysis of Labor Department Data by business journals revealed that “**U.S. financial sector loses 459,400 jobs in four years**” between 2008 and 2012.

We are investigating whether the upward trend of HFT played a role in this numbers.
Data & Methodology

- The data we use comes from two sources:
  - High frequency source (NASDAQ high frequency data)
  - Low frequency source (CRSP)
- Economic and labor data from Bureau of Labor statistics.
- We decompose HFT into its strategies identified by SEC (2010)
  - Passive market making (i.e. limit orders)
  - Arbitrage (i.e. benefiting from pricing inefficiencies)
  - Structural (i.e. exploiting market structure inefficiencies)
  - Directional (i.e. order anticipation and momentum ignition)
Mixed Data Sampling

- Since this study has mixed frequency data, we will use Mixed Data Sampling (MIDAS) regression introduced by Ghysels (2004).

- MIDAS is a class of time series models which allows the independent and dependant variables to be sampled at different frequencies.

- MIDAS is related to distributed lag models. Finite distributed lag model can be formalized as

  \[ Y_t = \beta_0 + B(L)X_t + \epsilon_t \]

- Simple MIDAS model according to Ghysels (2004) is

  \[ y_t = \beta_0 + \beta_1 B(L^{1/m};\theta)x_t^{(m)} + \epsilon_t^{(m)}. \]

Where

- \( B(L^{1/m};\theta) = \sum_{k=0}^{K} B(k;\theta)L^{k/m} \)
- \( L^{1/m} \) is the lag operator (captured by function \( B(L^{1/m};\theta) \) of few parameters summarized in vector \( \theta \).
- \( x_t^{(m)} \) is the number of lags
- \( \beta_1 \) is the overall impact of lagged \( x_t^{(m)} \) on \( y_t \)
Mixed Data Sampling

- **Parametrization of B(k;θ)**
  - Exponential Almon Lag
    \[
    B(k; \theta) = \frac{e^{\theta_1 k_1 + \cdots + \theta_Q k_Q}}{\sum_{k=1}^{K} e^{\theta_1 K_1 + \cdots + \theta_Q k_Q}}
    \]

- **In case of tick-by-tick data**
  - instead of using lag operator $L^{1/m}$, we can use operator $L^\tau$ where $\tau$ is real valued instead of rational number.
  - The Almon –type weight for the $\tau^{th}$ lag becomes
    \[
    b(k; \theta) = \frac{e^{\theta_1 k_1 + \cdots + \theta_Q k_Q}}{\sum_{k=1}^{K} e^{\theta_1 K_1 + \cdots + \theta_Q k_Q}}
    \]
Our study will be one of few studies which empirically apply MIDAS regression on tick-by-tick data.

We expect to produce a quantification of how each of the four HFT activity categories affect volatility and price discovery both in the short and long terms, whenever these effects are statistically significant, i.e. populating the table below:

<table>
<thead>
<tr>
<th>Pass</th>
<th>Short Term effect on</th>
<th>Long Term effect on</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price Discovery</td>
<td>Market Volatility</td>
</tr>
<tr>
<td></td>
<td>Price Discovery</td>
<td>Market Volatility</td>
</tr>
<tr>
<td>Passive Market Making</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arbitrage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Directional</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This data analysis will inform the policy side, by linking the dots from previous studies. By providing a coherent picture of HFT influence on market quality, it will allow cost-benefit analyses qualitatively and quantitatively, in which policymakers can assign value to liquidity and risk.
THANK YOU
References