Report, Views and Recommendations on "Modeling High-Frequency Trading Activity"

by Jean-Philippe Bouchaud, Paul Embrechts, Vladimir Filimonov, , Ramo Gençay, Larry Harris, Sergey Ivliev, Andrei Kirilenko, Albert S. Kyle, Guilherme Lamacié, Fabrizio Lillo, Richard Olsen, Christine Parlour, Alec Schmidt, Gilles Zumbach

September 1- 6, 2013, Banff International Research Station, Banff, Alberta, Canada
http://www.birs.ca/events/2013/5-day-workshops/13w5008

Overview

The workshop on high frequency trading assembled a unique group of participants combining expertise from theory to practice. The participants included practitioners with hands on experience working in hedge funds and organized exchanges, regulators wrestling with public policy issues, and theoretical and empirical researchers working in central banks and in academia. The presentations and discussions clearly showed how intricate the issue of high frequency trading is and how many factors determine its impact.

This document summarizes the discussions on the definition of high frequency trading, the effects of deregulation and technological progress on the emergence of high-frequency trading, policy & regulatory issues and recommendations. It concludes with a plea for more data and coordinated research.

What is high frequency trading?

While no universally accepted definition exists, the term high frequency trading (HFT) broadly refers to trading activity with position durations measured in microseconds, seconds, or minutes at the most.

HFT trading volume rapidly increased with the arrival of new technologies and regulatory opportunities to the markets. Starting in the 1970s, commissions were unfixed. Previously, high fixed commissions essentially functioned as a transaction tax, with the revenues being shared by the exchanges and the brokers. The deregulation of commissions increased competition and lowered transactions costs. The declining costs of computing power allowed exchanges and traders to adopt automated systems to implement new and faster trading strategies that transformed trading. In addition, the fragmentation of market centers increased the advantage to effective computing. Holding periods for many traders dropped as they implemented new strategies, many of which are designed to minimize risks.

One view of HFT is that it is simply a different technology for doing tasks that in earlier days humans once did on trading floors at exchanges such as the New York Stock Exchange or the Chicago Mercantile Exchange. For example, many of these traders simply tried to buy at the bid and sell at the offer to capture the bid-ask spread (or a fraction of it) as profit. This dealing strategy allowed them to make a nice living. With the introduction of electronic trading markets, the speed and accuracy with which computers can do these dealing trades forced human dealers out of the business. In this view, the computers simply do the same tasks that humans once did.

A second view of HFT goes beyond simply automating existing trading strategies. HFT also includes strategies never possible before such as high frequency quoting, low latency trading, algorithmic trading, statistical arbitrage, and others.

Although a well-accepted ecology and classification of HFT has not yet emerged, academics, industry practitioners and regulators generally understand that different types of high frequency trading can have very different footprints and widely different consequences for the market economy.

<u>How do deregulation and technology change market structure?</u>

The growth of high frequency trading is a consequence of deregulation in the US and in Europe. Regulatory changes converted comfortable monopolies or duopolies into fragmented markets, especially in the equity markets. High frequency trading emerged as a response to this fragmentation in the equity markets. For instance, technologically equipped fast trading outfits often engage in arbitrage strategies that involve trading quickly and simultaneously in multiple markets.

Increased market fragmentation undermined some of the benefits of centralized markets. For example, fragmentation reduces potential economies of scale. These lost economies, and the costs of arbitrage, may be socially wasteful.

At the same time, the increased competition and the introduction of new automated trading systems employed by exchanges and traders resulted in significantly smaller costs of transacting on exchanges. These reduced costs do not necessarily imply that high frequency traders create positive externalities, though many people think so.

Regulations also decreased tick sizes. In the US equity markets, the tick decreased from one-eighth dollar to a penny, a factor of twelve and a half. The decrease in tick size had a profound effect on quotation activity. To see why, image that the underlying value process follows a Brownian motion, and those traders update their quotes every time value crosses a discrete price. Under these assumptions, the number of quote updates is proportional to the square of the tick size. Thus a change in tick size from 12.5 cents to one cent will increase quotation active by more than 150 times.

Overall quotation activity also increases with increases in the number of trading venues. An increase in trading venues from one to ten would seem to introduce another factor of ten raising the overall factor to fifteen hundred. But since traders also update quotes in response to quote updates at other venues, quotation activity actually increases more with each additional venue.

The huge number of quotations requires large investments in the technologies required to receive, store, retrieve, disseminate, and monitor quotes. The problems associated with handling so much data in real time have led to glitches in automated trading systems used by markets and traders. These problems could be mitigated by consolidating order flow to a single exchange, but at the cost of losing the competition among exchanges.

Policy and Regulatory Proposals:

Markets face price distortions resulting from both run-away human traders and run-away computer algorithms. The best regulator of foolish behavior is deep appreciation and substantial ownership of the consequences. Knight trading distorted prices[1] due to run-away computer algorithms related to high frequency trading, lost hundreds of millions of dollars, and nearly went bankrupt. Société Générale distorted prices as the result of a human rogue trader exploiting low technology. Unwinding the rogue trades cost the firm billions of Euros. The Flash Crash resulted from humans choosing an execution algorithm that dumped huge quantities on the market over a short period of time, resulting in transitory price dislocations which cost the firm more than $100 million in incremental transactions costs. In all these cases, ownership of the consequences of their behavior was costly for the firms responsible.

Several regulatory proposals might improve market performance while undermining the profitability of high frequency trading. Small fees on messages reduce the advantages high frequency traders gain over other traders as a result of an ability to update quotes more quickly than other traders. Holding batch auctions, say every one second, would prevent high frequency traders from exploiting millisecond level information processing speed advantages during the one second between batch auctions. Slowing down trading at moments which high frequency traders would profit from the inability of other traders to provide liquidity would undermine the profitability of high frequency trading but might make prices more resilient.

Imposing an obligation on high frequency traders to act as market makers by providing two-sided quotes at all times ---including when they would like to withdraw bids and offers --- has been proposed as a way to make liquidity provision more continuous during periods of market stress. The NYSE had many years of experience with such policies but never got them to work satisfactorily. Strict rules might bankrupt high frequency traders during periods of extreme stress. Such rules would benefit the high frequency trading strategies which consume liquidity by placing executable orders which hit bids and lift offers provided by other market makers. Current rules punish brokers for making small mistakes, discouraging behavior designed to help customers at low cost.

---

[1] We should not attribute Knight Capital (KC) failure entirely to misbehaving algorithmic activity. According to the SEC report (http://www.sec.gov/litigation/admin/2013/34-70694.pdf), the major problem was due to (a) a human factor on the deployment of the new code for the trading system (the operator simply forgot to replace the old code at one of 8 servers) and (b) the absence of a desk (and standard procedures) responsible for such critical situations. In short, on August 1, 2012, there were two conflicting trading systems operating that dumped the price, and nobody in KC knew how to react to this situation.

Speed of Light, One-Second Batch Auctions, Fragmented versus Consolidated Order Flow

Technology has improved so much that the speed of light is quickly becoming a binding constraint on trading speed. Eventually, high frequency trading technology will become so cheap that all traders will have access to high frequency technology.

Those technologies will not always be equal, however. Traders who are faster than other traders will always have an advantage. The incentives to continue investing in high frequency technologies thus will always remain absent from regulatory interventions.

Markets will make faster or better access to trading opportunities available to some traders, like today's high frequency traders, and not to others, only if regulation and market organization allows it. To the extent that exchanges and the vendors of high speed communication services have market power, they will charge high frequency traders substantial fees that will extract much of the profits associated with high frequency trading.

The competition among high frequency traders for speed is an arms race in which only a few traders will survive. The surviving traders will generate greater revenues, which they will have to share with the exchanges and communication service vendors. The additional revenues from reduced competition among the high frequency traders will hurt buy-side traders. Regulators can substantially slow the arms race by introducing one-second batch auctions or by randomly reordering the processing of orders within sliding time windows of no more than 10 milliseconds.

Economies of scale and network externalities favor centralizing order flow processing at one exchange. Unsophisticated retail traders are protected by participating in a centralized exchange which consolidates order flow. Each individual trader wants to observe as much information about what other traders are doing while sharing as little information as possible about his own trading. These incentives to look but not be seen induce sophisticated traders to participate in dark pools. Although one goal of dark pools is to hide trading from high frequency traders, dark pools also create opportunities for high frequency traders to arbitrage one dark pool against another. Having fewer trading venues reduces by orders of magnitude the amount of data high frequency trading algorithms need to process. Thus, high frequency traders are likely to benefit from market fragmentation and are likely to be less profitable in a consolidated market in which all traders have access to state-of-the-art technology.

Did high frequency traders cause the flash crash?[2]

On May 6, 2010, in the course of about 36 minutes starting at 2:32pm ET, U.S. financial markets experienced one of the most turbulent periods in their history. Broad stock market indices – the S&P 500, the Nasdaq 100, and the Russell 2000 – collapsed and rebounded with extraordinary

[2] This and next subsections contain text taken verbatim from a study discussed at the workshop.

velocity. The Dow Jones Industrial Average (DJIA) experienced the biggest intraday point decline in its entire history. Stock index futures, options, and exchange-traded funds, as well as individual stocks experienced extraordinary price volatility often accompanied by spikes in trading volume. Because these dramatic events happened so quickly, the events of May 6, 2010, have become known as the "Flash Crash".

The event spread through the entire U.S. financial market system with extraordinary velocity and left a broad universe of market participants from professional traders with decades of experience to small retail investors with a realization that something was terribly wrong inside the multitude of automated markets. Calls for stricter regulation or even an outright ban of high frequency trading quickly followed.

On September 30, 2010, the staffs of the Commodity Futures Trading Commission (CFTC) and Securities and Exchange Commission (SEC) issued a report on the events of May 6, 2010. The 104-page report described how an automated execution program to sell 75,000 contracts of the E-Mini S&P 500 futures, algorithmic trading activity, and obscure order submission practices all conspired to create the Flash Crash.

The CFTC-SEC report was based in part on the empirical analysis released (concurrently with the CFTC-SEC report) as a working paper and discussed at the workshop. The empirical analysis based on regulatory (audit-trail) data shows that HFT did not cause the Flash Crash, but contributed to extraordinary market volatility experienced on May 6, 2010. The analysis also shows how high frequency trading contributes to flash-crash-type events by exploiting short-lived imbalances in market conditions.

Flash-crash-type events temporarily shake the confidence of some market participants, but probably have little impact on the ability of financial markets to allocate resources and risks. These events though raise a broader set of questions about the optimal market structure of automated markets.

Flash-crash events due to order flow imbalances may be ameliorated by slowing down trading. Slowing down trading gives non-HFTs time to place orders to stabilize the market in the presence of order flow imbalances. Alternatively, order imbalances can be resolved by altering the queuing algorithm and give preference to market participants quoting two-way prices; this would penalize traders that increase the order imbalance.

Bats or Birds: Ecology, Classification and Liquidity Provision

Growing empirical evidence suggests that contrary to some earlier studies, HFTs do not always act as liquidity providers. Broadly characterized, HFTs generally supply and demand nearly the same amount of liquidity. Advances in technology and infrastructure have altered the cost-benefit balance in favor of the most technologically advanced financial intermediaries ("bats") with the smallest overhead per market – the very definition of high frequency traders. Because traders can deploy advanced trading technology with little alteration across many automated markets, the cost of providing intermediation services per market has fallen drastically. As a result, the supply of intermediation services provided by the HFTs has skyrocketed. At the same time, the benefits of intermediation accrue disproportionally to those who possess the technology to take advantage of it. As a result, HFTs also have become the main beneficiaries of intermediation, using it not only to lower their adverse selection costs, but also to take advantage of the customers who dislike adverse selection, but do not have the technology to be able to trade as quickly as they would like to ("birds").

Consequently, the criteria of many definitions of HFT, e.g. U.S. Securities and Exchange Commission (SEC, www.sec.gov) and International Organization of Securities Commissions (IOSCO, www.iosco.org) may not be empirically valid.

How has high frequency trading changed indicators of market quality?

High frequency trading substantially reduced average bid/ask spreads. Its impact on spreads in different market scenarios, such as in response to news releases, is less well known.

High frequency trading has also altered patterns of market price volatility. Studies show high frequency trading sometimes contributes to accelerating price cascades, for example during the Flash Crash. Researchers are actively studying the feedback loops between high frequency trading and price dynamics to determine the circumstances under which high frequency trading accelerates, slows down or even reverses price trends.

To the extent that HFT exploits very short term modest predictability of asset prices, it tends to make asset prices adjust to information quickly. If short-term volatility otherwise would be less than long-term volatility, this HFT will tend to increase short term volatility to make it more consistent with long-term volatility. HFTs hold sufficiently small inventories that they do not prevent large price dislocations resulting from gigantic order flow imbalances. Indeed, HFT can contribute to accelerating price cascades.

However, for the ongoing and future research, when more specific market quality indicators are considered, there is eventually a necessity to adopt a more refined categorization within the class of HFTs itself. An HFT may behave quite differently from another, and that may depend on the properties of each market, e.g. liquidity aspects of a market. In a more granular analysis, one may expect to find quite aggressive and very passive HFTs. As such, different HFTs may have different consequences to the markets. That is why the question "good or evil?" is misleading, since it presumes a generalization of a property which is not necessarily generic.

After understanding the different types of HFTs and their externalities, it will be important to research how specific market conditions attract or repel each type of HFT. These market conditions include not only liquidity, price dynamics, risks, the mix of investors, and other microstructure properties, but also regulation and, last but not least, the detailed technological set-up.

This final element, summed to the importance HFTs have on the markets both in number of trades and in number of orders, yields a special and fatal property currently realized: the growing sensitivity of market quality (and the ensuing financial properties of traded assets) to the technological conditions of the marketplace, including the specific characteristics and robustness of machines, systems and networks.

The pricing model of the exchange, (normal or "inverted (taker/maker)" with rebates for liquidity taking and costs for liquidity provision), and type of matching engine (first-in-first-served, pro-rata or mixed) are extremely important factors that determine both high-frequency quotation activity and susceptibility to malicious HFT strategies. There have been developments in the industry that (implicitly) address some of the problems with HFT. For instance, some trading platforms have switched trading of some instruments from maker/taker to taker/maker pricing (in the latter, takers rather than makers are paid the rebates) and is an active proponent of abandoning the maker/taker pricing, which certainly motivates HFT to engage in market making. In FX, randomization of batched orders (each 1 - 3 milliseconds) has been rolled out.

While it is clear that HFT has decreased the bid-ask spread, it is important to investigate whether the advent of HFT has increased the market impact cost, due to the possible front-running activities. For medium-sized and large investors, market impact is the most important transaction cost, and studies on indicator of market quality should consider this important aspect.

The "quality" of a market is difficult to assess. There are a few metrics, such as bid/ask spread, short term volatility, or the amount of liquidity in the queues, but they provide for a very partial measure of the properties of a market, in particular with respect to its dynamics and resilience. The major difficulty is that the largest parts of the orders are latent. Without proper metrics, one cannot make the objective of evaluation of what constitutes “good or "bad" market, or of a trading practice.

Financial markets appear statistically very similar, across geographical areas and financial instruments. In particular, tail events occur with similar frequencies. In view of this regularity, it is not trivial that the market rules can be modified to obtain a more resilient market, with less extreme shocks.

What amount of intermediation is correct? Clearly, HFT increases intermediation. Is this good, or too much?

Data for Academic Research: A Necessity for a More Resilient Market Design

Trader-specific trade and quote data must be made available for academic research. Regulators and trading venues must solve issues of data confidentiality so that they can provide comprehensive samples of market data to academic researchers. Regulators should mandate that exchanges release historical properly anonymized and stamped data samples to academics to stimulate research projects that would eventually lead to a more resilient market design.

Empirical studies are clearly needed on the how various high frequency trading strategies are based upon and affect trade volume, number of messages, relative net inventory change at the end of the day, number of reversions (crossing zero/mean), participation at the top of the book, ratio of active versus passive orders, maximum intraday position to volume traded ratio, etc. Common classification and understanding of HFT is crucial for a comparability of various studies. Doing this type of analysis on the pre-HFT data also could help answer the "birds" versus "bats" question.

The market data explosion has made it difficult to analyze high frequency "big" data. Some markets now have securities with more than 100 million messages per instrument per day. Academic institutions may choose to collaborate and join efforts to create common facilities to maintain market data and high-performance computing facilities, as well as matching engines that would allow one to replay the market and/or experiment with synthetic flows. Such distributed high-performance computing facilities already exist in other sciences and similar models can be pursued.

Methodology for the analysis and modeling of HFT data

During the meeting, several discussions took place on the topic of mathematical tools for modeling HFT data. For instance, an evening session was devoted to the theory and applications of Hawkes processes, self-exciting processes. These processes allow for a versatile modeling of bursts of activity across different time frequencies, and these are both for one-dimensional as well as more-dimensional data. Progress was reported both on the probabilistic modeling, for instance related to the notion of stability, as well as on statistical estimation, both in the frequency as well as time domain. Further discussions involved the stochastic generation of heavy-tailed, power-tail like models. They play a crucial role for the modeling of HFT return data distributions. Finally, operational risk modeling in view of HFT was reviewed. The main stress was put at the level of risk aggregation, model uncertainty and especially dependence uncertainty.

Thank you to all participants for your contributions at the workshop and in this report:

Agliardi, Rossella
Ait-Sahalia, Yacine
Almgren, Robert
Andersen, Torben
Ao, Han
Bae, Kyounghun
Bell, Peter
Bouchaud, Jean-Philippe
Brogaard, Jonathan
Chaboud, Alain
Djurdjevic, Deana
Embrechts, Paul
Filimonov, Vladimir
Foucault, Thierry
Garriott, Corey
Gencay, Ramo
Gradojevic, Nikola
Gray, Kyle
Harris, Larry
Hendershott, Terrence
Ivliev, Sergey
Jones, Robert
Kasa, Ken
Kirchner, Matthias
Kirilenko, Andrei
Kyle, Albert S.
Lamacie, Guilherme
Li, Wei
Lillo, Fabrizio
Mahmoodzadeh, Soheil
Olsen, Richard
Parlour, Christine
Ping, Li
Schmidt, Alec
Stanley, Eugene
Sutter, Barbara Caroline
Tsang, Edward
Tseng, Michael
Vatter, Thibault
Xu, Jiangmin
Xue, Yi
Yu, Xiao (Christy)
Zhang, Kerry
Zumbach, Gilles

http://www.birs.ca/events/2013/5-day-workshops/13w5008