Government Office for Science

Foresight

Computer-based trading and market abuse

Driver Review DR20

Foresight, Government Office for Science

Contents

Computer-based trading and market abuse

Sylvain Friederich and Richard Payne[1]

15 October 2012

This review has been commissioned as part of the UK Government's Foresight Project, The Future of Computer Trading in Financial Markets. The views expressed do not represent the policy of any Government or organisation.

[1] School of Economics, Finance and Management, University of Bristol and Cass Business School, respectively.
We thank three anonymous reviewers for their comments. All remaining errors are ours.

Executive summary

Trading is now conducted on a scale of microseconds by some market participants. Does extremely reduced latency increase the likelihood of market abuse – in particular market manipulation?

We argue that:

- It seems that some classes of participants perceive abuse as more widespread in today's markets, where HFTs may dominate liquidity supply.
    - This perception has in itself the potential to affect the behaviour of less-informed agents, particularly liquidity suppliers, thus impacting market outcomes negatively.
- The concern that high-frequency traders use a speed advantage over other agents to implement new abusive strategies generates much speculation but other issues may be equally important.
    - In particular, the growth of HFT may have altered the trading environment in ways that render some forms of abuse easier to conduct than in the past.

As a result, we suggest that:

- Perception must be confirmed or corrected through the production of hard evidence on the link between HFT and abuse.
- Significant investment is likely to be required for regulators to surmount the surveillance issues that HFT has created.

> HF traders offer little benefit to the financial system, but they do not do much harm.
>
> (Lex, *Financial Times*, 12 July 2011)

The view above is, to say the least, not a typical one. Much investor and regulator suspicion surrounds high-frequency trading (HFT). In Europe, it has been placed near the top of the regulatory agenda with the release of a Consultation Paper entirely devoted to it (ESMA 2011).[2] In the US, the SEC spent over 20 pages of its "Concept Release on Equity Market Structure" detailing its concerns (SEC 2010) and the global committee of regulators (IOSCO 2011) just released its own report largely dedicated to HFT.

[2] This is a striking change of emphasis. Only a few months before, HFT did not register much in the EU Commission's consultation leading to the update of the main piece of legislation that regulates trading across the EU, the Markets in Financial Instruments Directive (see European Commission 2010, where HFT only receives a passing mention).

Chief among the concerns are the effects that HFT may have on liquidity, volatility and the probability of flash crashes. Additionally, a possible link between HFT and market abuse is frequently alluded to – for example in each of the three reports mentioned above. This link is the focus of the current paper. Taking market abuse to broadly comprise *insider trading* and *market manipulation,* we discuss three, not mutually exclusive ways in which HFT may matter to market abuse:

- Low latency may have increased the incidence of market abuse because speed allows HF traders to take advantage of slower market participants in abusive ways.

- It could also be that the growth of HFT has changed the trading environment in ways that facilitate market abuse or have increased the perceived extent of abuse. This covers situations different from the 'fast versus slow agent' confrontation above, where for example HF traders prey on each other and slower traders may also find it easier to conduct abuse.

- Other market developments concomitant with the growth in HFT but not only or at all brought about by HFT growth may also have contributed to an increase in the perceived prevalence of "unfairness" or downright abuse. These are confounding factors that are part of the new trading environment and may be hard to disentangle from HFT empirically, even though they are conceptually very distinct.

The first view implies that lower trading latency allows HF traders to run rings around other traders. Along the second and third views, the trading landscape has changed in ways that may generate more abuse or make abuse perceived as more widespread, but HFTs are no more likely than other traders to engage in abuse.

We argue that even though evidence supporting the first view is currently lacking, the perception of a higher incidence of abuse should be dismissed, *whether or not this perception is divorced from reality*: a thread running through this report will be that the subjective estimate of abuse that market participants form affects their behaviour and has the potential to directly impact market outcomes.

Different regulatory implications may follow. If the key concern is that HFT allows some participants to abuse the market in ways that are not currently unlawful, then regulatory definitions of what constitutes abusive practices may have to be modified and new methods to detect such practices will need to be developed. If HFT makes old forms of market abuse more widespread or socially more costly than before, then exchanges and their regulators must improve their ability to root out such practices. If perception is mistaken, then hard evidence should be produced to document the reality of abuse.

This debate is part of a broader one on HFT and redistribution. Extant academic studies of HFT tend to find that it has a positive effect on average liquidity, at least based on simple measures like bid-ask spreads. However, very little is known about the potential redistributive effects of computer-driven trading: who gains and who loses? The focus of this report is whether HFT may have affected the scale of redistribution generated from unlawful behaviour, and whether practices which might previously have been considered abusive are now accepted trading behaviour.

As background to this, the suspicion that most market abuse goes unpunished is widely held. The UK market regulator recently declared: "Our benchmark should seek to have a market that participants really believe to be clean and fair," (…) "And, as a general test, I think that if you were to ask the market participants, they would share my view that there is too much market abuse."[3] However, abusive trading behaviour is easy enough to define in very general terms but hard to identify. Regulations against abuse are rarely enforced because it so difficult to demonstrate in a court of law that the *intention* of a trader was to manipulate prices or to profit from privately held information.[4] As a result, there is scant empirical evidence on the true extent of abuse in financial markets, and none to our knowledge on abuse specifically conducted using computer-based strategies. (In turn, the difficulties involved in gathering evidence demonstrating a clear intent to abuse the market may also explain the propensity of some regulators to proceed in what seems like an evidence vacuum to academics or industry.)

Therefore this report will be necessarily speculative in nature. We will use economic theory, regulatory literature and qualitative evidence obtained in interviews with slow and fast traders. Further empirical investigation in this area is both feasible and desirable.

We start below with definitions of market abuse and of HFT. We then discuss how they may be related.

1. What is market abuse?

Economists and regulators identify two broad types of abusive market behaviour: *insider trading* and *market manipulation*. The key difference between them is that the former profits from price movements that are expected to happen, while the latter attempts to engineer price movements to profit from (equivalently, one is based on privileged information and the other typically is not).

Insider dealing involves trading on information that is not publically available at the time of the trade but is expected to move security prices in the future. A regulatory definition of insider trading would typically add two requirements:

- First, the information traded upon must be expected to move security prices *in the near future.* Corporate insiders are allowed take a long-term view of the prospects of their firm -- indeed, this forms the basis of many executive compensation schemes -- but are not supposed to trade around the release of immediately "price-sensitive" information.

- Second and crucially to regulators, trading is based on information that was obtained at no cost. The distinction between an 'insider' and any profit-seeking investor resides in this point: insiders did not have to expend resources accessing and processing information but became privy to high quality information in the course of their normal business activities.

[3] Hector Sants, cited by the BBC, 14 March 2010, http://news.bbc.co.uk/go/pr/fr/-/1/hi/business/8566904.stm

[4] In particular, only successful instances of abuse can usually be prosecuted. Unlucky or incompetent manipulators are very likely to escape scot-free.

Market manipulation broadly implies an ability to move prices temporarily to one's advantage. Before prices return to equilibrium, the trader unwinds their position at a profit. Prices can be manipulated in two main ways:

- First, through credibly *appearing* informed. This refers to techniques such as bluffing, which may be conducted using various channels (e.g. internet-based investor discussion groups).[5]

- Second, and mostly, through trading. This type of manipulation requires the ability to exert temporary *price pressure,* a reflection of the fact that markets are less than perfectly liquid. (In a fully liquid market, prices only move in response to the release of new information relevant to future security payoffs.)

Beyond very broad definitions though, the legal and the economic approach to abuse differ substantially in their emphasis. We consider them in turn.

2. Regulatory and economic views of market abuse

2.1. Legal/regulatory definitions of market abuse

In Europe, the 2003 EU Market Abuse Directive (MAD) specifies that:

> Market abuse may arise in circumstances where investors have been unreasonably disadvantaged, directly or indirectly, by others who:

- Have used information which is not publicly available (insider dealing);

- Have distorted the price-setting mechanism of financial instruments;

- Have disseminated false or misleading information.

> This type of conduct can undermine the general principle that all investors must be placed on an equal footing.

Translating this Directive into local law, the UK market regulator (Financial Services Authority or FSA) lists the "types of behaviour [that] can amount to market abuse":

- Insider dealing, including:

- "Improper disclosure" of inside information to a third party and

- "Misuse of information" that can affect the share price,

- Manipulating "transactions" or "devices"

[5] This concern is at the heart of the 2011 bans on short-selling on European exchanges. It was hypothesized that short-sellers were establishing positions in financial firms before spreading false, negative rumours about the financial health of those firms. If prices drop in response to the rumours then the short-sellers can close their positions at a profit. There is no direct evidence for this type of manipulative activity occurring.

- Dissemination of “false or misleading information”

- “Distortion and misleading behaviour”, defined as “behaviour that gives a false or misleading impression of either the supply of, or demand for, an investment; or behaviour that otherwise distorts the market in an investment.”

Legal definitions of market abuse largely reflect the view that it should be prevented because it is unfair to someone.[6]

There is not much doubt that company directors and perhaps other constituencies (e.g. large shareholders and fund managers) are regularly aware of information that share prices do not yet reflect: academic studies of *legal* insider trading (conducted by company directors when they are allowed to do so) indicate that these trades have significant predictive power for long-term returns.[7] Under insider trading, the unfairness arises not so much because the trades of insiders are likely to inflict a trading loss to their counterparties, but because this information was ‘privileged’ i.e. received by virtue of the insider’s functions and not by the resources that they expended seeking it. As a result of these views, the trades of insiders are tightly regulated on many markets and therefore it is arguably the case that the actual prevalence of insider trading is largely determined not by the amount of information insiders hold but by the extent to which they feel able to exploit it. This is itself a function of both how strict and how tightly enforced local regulations are.

2.2. Economic approaches to market abuse

Economic theory does not tend to focus on the distinctions that matter to regulators between short versus long term price changes or the lack of cost incurred to become informed. Neither do economists enter into definitions of fairness in trading outcomes that they feel are not within their domain – they prefer the more neutral term of ‘redistribution’. An insider is just an agent holding very precise information, such that interaction with them is likely to generate a trading loss. The focus of economists in the area of abuse is on two directly related issues:

First, they are concerned with identifying the *market-wide* implications of abuse: what is the effect of abusive behaviour on variables such as activity, liquidity, pricing efficiency and volatility?

Second, whereas regulators worry about identifying the actual victims of abuse and preventing further instances of it, economists are concerned with how the *perceived* likelihood of abuse, as distinct from its actual occurrence, may affect market outcomes.

- Just like the true value of an equity share, the true extent of abuse will never be known. Investors can only attempt to infer it.

- Because trading is inherently strategic, the subjective probability that various classes of agents, especially those supplying liquidity, attach to the incidence of abuse, will alter their

[6] Note that the definition of fairness relevant to this context is open to debate.

[7] See Seyhun, N. (1998), *Investment Intelligence from Insider Trading*, MIT Press.

behaviour. Therefore the *perception* of abuse is capable of reducing activity and damaging liquidity. It matters even if it is for a time disconnected from the *reality* of abuse.

The box below illustrates this difference in emphasis between the regulatory and the economic approaches to market abuse, specifically on the topic of insider trading.

Finance theory and insider trading

Regulators and the investing public tend to oppose insider trading on grounds that those trading with insiders are likely to lose money. Financial economists would emphasise two sets of implications:

First, insider trading may improve pricing efficiency.

- If insiders who are privy to price-sensitive information before it is made public were allowed to trade on it without restriction, their trading should cause prices to reflect information faster and markets may perform their allocational role better, other things equal. This improvement in efficiency thus generates a benefit for society, although it comes at the expense of some traders losing out from trading with insiders.

Second though, insider trading may damage market liquidity. In the absence of insider trading regulation, the increased likelihood of trading against someone better informed than they are would cause less informed market participants to modify their behaviour. Their change in behaviour would likely depend on the subjective estimate they would form of the probability of losing out to better-informed traders:

- Under moderate informational asymmetries, less informed agents might avoid exposing their trading intentions to the market by selecting pre-trade opaque trading venues (OTC, dark pools). Market transparency deteriorates and so may liquidity on transparent venues.

- Under more severe informational asymmetries, uninformed investors might reduce their participation in the equity market, causing a vicious circle of illiquidity that could become, in extreme circumstances, a market breakdown caused by *adverse selection* (where only informed agents are left to trade).

- Under less extreme but plausible levels of informational asymmetries, some uninformed traders may exit the market leading to declines in average liquidity and perhaps also pricing efficiency, due to the decrease in operational efficiency. (The operational efficiency of a financial market, as opposed to its informational or pricing efficiency, refers to the level of transaction costs incurred in trading. These two concepts of efficiency are distinct but directly related: in a market or an asset with high trading costs, it may not be profitable to implement an arbitrage strategy that will correct a mispricing,

and the mispricing may persist.)

This illustrates the two points made above: the economic approach to abuse is concerned with the impact that the *perceived* likelihood of informed trading by less-informed agents has on *market-wide* outcomes. Their reaction may range from concealing part of their trading intentions to withdrawing from the market altogether. The reasoning is illustrated here in the context of insider trading specifically, but it applies to any form of abuse.

Kyle and Viswanathan (2008) take economic reasoning to its logical conclusion by arguing that trading behaviour should *only* be branded manipulative and be made illegal if it is *socially* harmful to the economy. Along this view, manipulation must demonstrate the "violator's intent (…) to pursue a scheme that undermines economic efficiency *both* by making prices less accurate as signals for efficient resource allocation *and* by making markets less liquid for risk transfer." In other words, price manipulation should only be made illegal if it at once prevents markets from fulfilling their role in guiding productive investment and increases the costs of trading financial assets.[8]

Regulators are also concerned with market outcomes of course. For example, they differentiate between market abuse and individual abuse such as securities fraud, where a clearly defined and limited set of investors was abused. Front-running is a case in point.[9] But whilst they recognize that abuse undermines confidence in the market at large, regulatory action cannot rest on attempts to demonstrate that market-wide liquidity was degraded by abusive behaviour. Prosecution requires identification of one or several victims.

For practical purposes, the regulatory and economic approaches are not inconsistent but they have very different emphasis.

We know very little about either side of the argument: how many victims market abuse really has and what the true market-wide effects are of the perceived likelihood of losing out to insiders or manipulators. To illustrate, the average number of "enforcement actions" taken every year by the SEC over the past ten years is 50 against insider trading and 39 against market manipulation.[10] These are not large numbers for what is by a long way the largest equity market in the world. Data from the UK FSA are more disaggregated and not directly comparable but the total number of actions represents perhaps half of the above (See FSA Annual Report 2010-11, pp. 59-60).

---

[8] This definition seems to leave open the possibility that insider trading could be allowed if it made prices more efficient. The sole focus of Kyle and Viswanathan in that paper is on manipulation though.

[9] In this respect, the UK regulator thus considered in the past that "*The Code is aimed at preventing conduct amounting to market abuse and FSA's view is that front-running customer orders is a customer abuse, not a market abuse. Clearly, the FSA does not condone front-running and while the Code does not restrict the use of order-flow information, the FSA Handbook will contain appropriate rules preventing such customer abuses. It is likely that these will appear in the Conduct of Business Sourcebook rather than the Market Conduct Sourcebook*." (FSA Feedback Statement on responses to Consultation Paper 10: Market Abuse, 1999).

[10] These "Enforcement Actions" may themselves entail a range of actions that greatly vary in their severity and implications, from Federal Court actions down to SEC 'opinions'. (Abuse may be punished in other ways: exchanges, as Self-Regulatory Organisations, may also take disciplinary action against abuse although they seldom do. Individuals may also invoke Federal and State laws if they feel they have been defrauded.)

These definitions illustrate the difference of emphasis between regulators and economists, and also how difficult it is to define what insiders or manipulators do beyond broad generalizations.

3. What is high-frequency trading?

To understand the ways in which HFT may increase the likelihood of abuse, we need to define HFT and what is new about it. The upshot of this section has similar flavour to that of the previous one: definitional problems abound. The US SEC (2010, p. 45) notes, "The term [HFT] is new and not yet clearly defined". From the name, one would suspect that HFT must involve in some sense 'fast' order submission and revision and, because it is computer-driven, extensive automation of trading processes. Beyond this, HFT is also typically taken to entail relatively small investment positions that are held for short periods, and trading that is conducted in a principal ("proprietary") capacity. But this can still mean all things to all people.[11]

A participant gives this definition:

> [HFT is] any trading that's faster than humanly possible. Any kind of trading that a human can't execute anymore, that's what I would call high frequency trading. Then I would make a distinction between high frequency trading and ultra-high frequency trading. High frequency trading would be something you can do in software, whereas ultra-high frequency trading is something you can try to do in software, but you are probably better off doing it in hardware. (Interview with Peter Van Kleef, High Frequency Trading Review, Sept 2010.)

This introduces two questions to which we come back later: should HFT be defined in terms of absolute or relative latency, and how significant are the speed differences within the class of HFT players?

For want of an agreed definition, it is at least useful to see what HFT is not. Indeed, part of the reason why debates surrounding HFT seem so divisive is that most observers find it difficult to get a handle on what high frequency traders actually 'do' beyond broad definitions. The other side of the coin is the constant complaint of HF traders themselves that what they do is misunderstood and misrepresented. This is because HFT is not a clearly delineated set of trading strategies and, for the most part, nor can these strategies only be conducted at a high frequency. One HF trader puts it this way:

> HFT is not a strategy; it's a means of execution. It's not a trading strategy, it's a business strategy. It's how you want to go about doing business, how you want to execute your trades and how you want to go about making money, but it is not in and of itself a trading strategy. That would be like saying, "that pit trading strategy" or "that long term investing strategy that all those mutual funds run". (Interview with Peter Nabicht, The High Frequency Trading Review, 7th July 2011)

Perhaps adding to the regulatory unease, HFT is not a standard feature of trading system design that is under the direct control of regulators or even exchanges, unlike for example tick sizes or the degree of trading transparency. It is a market-led development. Exchanges have

[11] For a more complete taxonomy of HF trading styles, see Friederich and Payne (2011a).

certainly done everything they could to accommodate it by offering faster mainframes, co-location services, smaller tick sizes, and by designing fee structures that are attractive to the HFT community. But they and other electronic venues were reacting to the demands of a substantial part of the industry and to the implications for their own business models, they did not engineer HFT. Regulators themselves are unsure of how they would go about curtailing HFT, assuming that this was their policy goal.

4. Why does HFT matter?

The short answer to this question is that we don't really know yet. Both academics and regulators were forced to think about HFT when it became so prevalent that it was impossible to ignore. The few microstructure models that explicitly address the latency issues involved in HFT and academics are very new and academics are lacking a consensus theoretical framework that would help them think more rigorously about how it might affect liquidity, volatility or systemic risk. In the box on the next page, we review some of the recent theoretical pieces which present arguments both for and against HFT in terms of efficiency and liquidity.

One approach is to ask what is new about HFT. The *roles* performed by computer-driven traders are not new. They are broadly of three types:

Order execution. This is the traditional function of the broker in any market. This order flow is generated in an agency, not proprietary, capacity and in an automated way but where the lowest latency is not typically paramount. This activity is largely the preserve of large "sell side" firms.

HFT and latency in theory

Several new pieces of finance theory allow for speed differentials between market participants and thus have the potential to speak to HFT issues. Below we review a few of these contributions.

Gerig and Michayluk (2010) embed an automated market-maker in a model based on Glosten and Milgrom (1985). A key advantage of the automated market-maker (AMM) is that it is able to effectively trade in multiple securities while traditional market-makers cannot do so. The fact that the AMM can process information from related securities means that the prices it sets will, in general, be better than those from traditional sources. The model results in more efficient prices, lower trading costs for liquidity traders and higher volumes. Thus, in this framework, automated trading activity is far from 'abusive'. Martinez and Rosu (2011) have a similarly positive view of HFT activity. Focussing on HFT traders who use market orders, rather than HFT market makers, they build a model in which HFT traders use their speed advantages make markets more efficient by quickly bringing information to market. Again, this could not be described as abusive trading.

A more worrying conclusion is reached by Jarrow and Protter (2011). They argue that the combined activity of HFTs who trade independently but on correlated signals may create mis-pricings in securities that are damaging to the welfare of non-HFT traders and increase volatility. The detrimental effects of HFTs are generated by their latency advantage and the authors argue that, where possible, regulators should attempt to eliminate latency differentials between traders. Cohen and Szpruch (2011) build on the work of Jarrow and Protter to show how HFTs may profitably *front-run* slower traders and how this may

reduce market efficiency. They go on to analyse whether a Tobin Tax might limit the effects of HFTs on market quality.

Cartea and Penalva (2011) generalize the Grossman-Miller framework to include a class of HFTs who are non-fundamental traders and exploit a speed advantage to step in and intermediate trade between uninformed liquidity traders (e.g retail traders) and professional traders. Their key result is that the inter-position of HFTs between the liquidity and professional traders leads to liquidity traders executing at worse prices (as HFTs extract some rent from their speed advantage).

Hoffman (2011) incorporates speed differentials into the Foucault (1999) framework where the key advantage of speed is the fact that, after revelation of public information, fast traders' orders are less likely to be picked off than those of slow traders. Hoffman's model predicts that, in certain scenarios, the advantage enjoyed by HFTs leads to non-HFTs suffering higher execution costs and also that total social welfare may be reduced. Thus, from this perspective HFT activity may be damaging even if it is not actively manipulative/abusive. Jovanovic and Menkveld (2011) study a related framework and argue that the reduced likelihood of HFTs being picked off around public information announcements makes them more likely to make markets and can thus improve welfare. However, HFT activity can reduce welfare in this model too if subsequent to information release, HFTs can exploit slower traders who are not aware of the information update.

As the preceding summaries indicate, the effects of HFTs on market quality and their potential to reduce efficiency and/or liquidity are still matters of debate. Note, though, that the empirical evidence available so far on HFT and market quality is highly weighted in favour of HFT, and that most of the theoretical models reviewed above do not focus on abuse.

Statistical arbitrage mostly involves high-frequency dealing in a principal capacity. 'Stat arb' covers a wide set of strategies conducted across a variety of assets or markets. Simple examples involve arbitraging price discrepancies in the same assets traded on different venues, stocks that have highly correlated returns ('pairs trading'), cash assets and their derivatives or Exchange-Traded Funds and their component securities. Computer-driven traders can lay claim to conducting stat arb more efficiently than traditional human traders, because speed allows them to detect mispricings much more quickly, and because several legs of an arbitrage strategy can be executed more simultaneously in time, reducing the risk that traditional statistical arbitrageurs incur. This role is typically not controversial and is considered conducive to higher pricing efficiency.

Market-making is essentially conducted at high frequency and as principal. There is much controversy surrounding this role. Many HF traders are keen to call themselves "market-makers" but this is in many cases inappropriate because they have no formal commitment to stay in the market. As a result, they are regularly accused by some regulators and by the institutional buy-side of leaving the markets when the going gets tough. It would be more adequate to call this class of HFTs 'voluntary liquidity suppliers'. They would not be expected to behave as traditional, 'affirmative obligation' market–makers and less criticism would be heaped on them as a result.

It's mostly the latter two roles that are considered 'core' HFT territory today because the lowest latency is considered an advantage in performing these functions. What 'real' HFT market-making there is takes place in assets where exchanges have retained or reinstated formal market-making to ensure continuous two-sided liquidity. Market making in Exchange-Traded Funds in particular is dominated by HF firms, because most of these funds are thinly traded (beyond a handful of them that track the largest equity indices), and because market-making in these assets is particularly closely tied to implementation of statistical arbitrage.

Described in this way, nothing seems very new. By many aspects, HFT seems like a continuation of developments that have been affecting market structures since the 1980s:

- The gradual demise of formal market-makers in liquid stocks (whether competing as on London's SEAQ or on the Nasdaq, or monopolistic, such as the Specialists on the NYSE). Formal market-making in active names was terminated in London in 1997.

- The automation and computerization of trading processes (allowed by less market-making, more order-driven, electronic structures).

    - In 1987, automatically triggered program trades were blamed for exacerbating the market crash; Professional market-makers have used software to automate quote refreshing much before the term 'high-frequency' was used; In the same spirit, proprietary trading has been implemented using trading 'engines' (scripts) for over a decade.

- The decrease in trading latency, again largely the result of point 1 above. Attempting to revise one's quotes or to trade faster than other players is nothing new.

    - See, e.g. the "SOES bandits" controversy in the early 1990s where fast traders using an electronic execution system in small sizes inflicted trading losses on those Nasdaq market markets that did not update their quotes quickly enough.[12]
    - Using U.S. data from 2000, Busse and Green (2001) examined whether prices reacted to analyst recommendations given on television. They found that "(…) prices respond to reports within seconds of initial mention, with positive reports fully incorporated within one minute."[13]
    - Order submission very quickly followed by cancellation was described as 'fleeting orders' by Hasbrouck and Saar (2009). Using data from 1999 and 2004, they noted: "In contrast to the usual view [that limit order traders are patient liquidity suppliers], we find that over one-third of nonmarketable limit orders are cancelled within two seconds."[14]
    - Regarding flash orders, Harris and Namvar (2011) note that "the function of flash facilities within electronic trading systems is consistent with practices that always have been accepted in floor trading markets. (...) The "flash period" on the floor was typically 15 to 30

[12] Harris, J. and P. Schultz (1998), "The Trading Profits of SOES Bandits", *Journal of Financial Economics*.

[13] Busse, J. and T. C. Green (2002), "Market Efficiency in Real Time", *Journal of Financial Economics*.

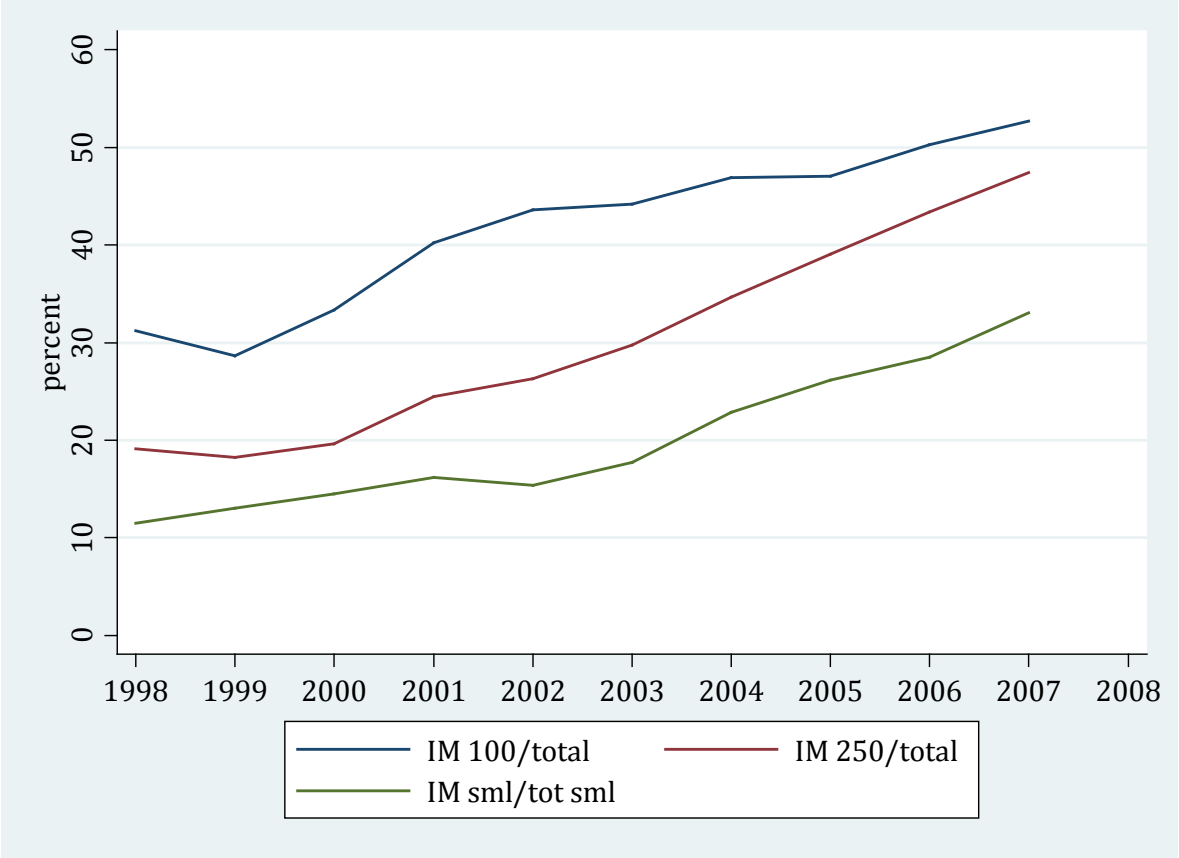
[14] Hasbrouck, J. and G. Saar (2009), "Technology and liquidity provision: The blurring of traditional definitions", *Journal of Financial Markets*.

seconds, but the notion of seeking additional liquidity from other participants in the exchange is essentially the same."

- Small investment positions and short inventory half-lives have been used since the 1990s by so-called "day traders".
    - The "SOES bandits" referred to above are an example. Using data from 1995, Harris and Schultz (1998) found that "Bandits usually hold a position for only a few minutes."
- The growth in 'worked' order execution and attendant decline in average trade size was spurred before HFT by the development of quantitative techniques in portfolio management, in particular formal models of dynamic market impact.[15] Increased investor sensitivity to the impact of trading costs on portfolio performance probably played a part and so may have the growth in benchmarking techniques (like VWAP, TWAP and other).
- The increase in proprietary trading has also been a long-term trend.
    - Figure 1 shows the proportion of total trading where both sides to the trade reported having dealt for own book on the LSE. The graph shows a continuous increase that much predates the growth of HF trading.

[15] A well-known reference is Almgren, R. and N. Chriss (2000), "Optimal execution of portfolio transactions", *Journal of Risk*.

Figure 1: The graph depicts the proportion of total trading that is classified "intra-market" by the LSE. Intra-market activity is defined as comprising trades where both sides reported having dealt in a Principal capacity (for own book). The Figure shows this proportion for components of the three main UK stock indices (FTSE 100, FTSE 250 and FTSE Small Cap) separately. (For data reasons, the graph ends in Q3/2007).



Clearly, the roles that HF algorithms play are not new, and the individual 'components' of their trading behaviour, such as an emphasis on speed or proprietary positions are very familiar too. They've been used before and for the same purposes, although HFTs may combine them in a new way and perform them with greater speed. It's tempting to conclude that HF Trading is just the same as before, only faster. This is the view of themselves that HFTs like to promote – they are not doing anything new, just performing old functions more efficiently.

4.1. The role of latency

Any discussion of HFT must attempt to clarify the role of latency. There are various ways that latency could be defined in a trading context, such as the time elapsing between a market event and the decision from a computer script to enter an order or revise a quote, or between the same decision and the moment the message implementing the order entry or quote revision reaches the Exchange's systems or as the time difference between an algorithm submitting an order and receiving an acknowledgement that the order has reached the market. The market 'event' could be the receipt of a piece of text containing news for a market-making script, or a change in the price of a related security for a stat arb script, or a change in order flow imbalance for an order execution algo.

A couple of decades ago, traders who could make and implement a trading decision in a matter of minutes could no doubt have been labeled "High-frequency" even though the term was not

used. In the 1990s, academics would refer to intraday data as 'high-frequency',[16] at a time when a few quote revisions per minute on average might have placed a stock somewhere in the top decile of liquidity. What constitutes 'high' frequency today? We are told that 250 milliseconds is an eternity to some players. When does a quote become a 'flash' quote? This suggests that an absolute definition of latency may be of limited relevance.

HFT is often thought of as in terms of a reduction in *average* latency. But perhaps more important are the ways in which the distribution of latency across market participants has been affected. HFT may matter to abuse as it changes the relative reaction times of various trading constituencies.

- To illustrate, consider claims by HFTs that they are able to supply liquidity better than traditional market–makers because they are quicker to update their quotes in response to news, which leaves them less vulnerable to adverse selection. This can only hold true if informed traders have not reduced their own trading latency by the same extent. In other words, it relies on an improvement in one's latency relative to that of others. In a world where the quoting and trading speed of all agents increases by the same factor, the only conceivable improvement becomes one of informational efficiency – prices adjust to news more quickly. (Whether it matters to efficient resource allocation that prices adjust in a matter of milliseconds rather than seconds or minutes is a matter of debate.)

This implies that a speed advantage at once gives rise to claims by HFT firms that they are better able to perform traditional roles *and* to accusations of an "uneven playing field" and sometimes market abuse on the part of slower agents.

A key consideration then becomes whether this speed advantage is within the reach of enough participants to make the market competitive or at least contestable.[17] What is the size of the barriers to entry to becoming an HFT player? Some observe that the profits of HF players are large but that only a handful of firms dominate order flow, which would be indicative of very high barriers to entry. Against that, one can note that if financial means were paramount, then the largest 'bulge bracket' investment banks that are both cash-rich and have always boasted of being glamorous enough to attract the best talent in the industry should dominate the field. Instead, many of the top HF players are firms that were only known to insiders a few years ago (Citadel, Optiver, Getco, Flow Traders, Susquehanna etc). This trend may be reinforced by recent regulation, e.g. the Volcker Rule, which has caused investment banks to scale down or exit proprietary trading activities that might have competed with HFT firms. An HF trader cited above gives the following view, consistent with HFT shops having upset the old order where major US investment banks ruled the roost:

> I would say the leaders in our market clearly are the small prop trading firms because they drive innovation. Usually you'll see whatever is the latest and greatest technology in the small trading shops first. Then it goes to bigger sell side firms, then it goes to advanced

[16] See, e.g., C.A.E. Goodhart and M. O'Hara (1997), "High frequency data in financial markets: Issues and applications", *Journal of Empirical Finance*.

[17] A market for a particular good or service is *contestable* if the barriers to entering or leaving this market and the sunk costs involved are low. As a result, even if only a few firms dominate that market, the *threat* that new competitors might enter it at any time serves to making the incumbents behave competitively.

> hedge funds, to normal hedge funds, and maybe a couple of years later it gets to the mainstream in terms of normal asset managers. So I think there's a natural progression in terms of people driving the technology, usually due to small firms needing to innovate, finding an edge in the market and their being nimble and small, so they can implement new technologies much quicker. If you have legacy infrastructure that cost $50 million, like a big investment bank might have, throwing all that out and buying all the latest and greatest kit is not going to be an option. But if you're a small firm with just 50 guys and some new kit that costs just a hundredth of the price, then you can easily buy that and implement it immediately.
>
> What then happens, the bigger firms see that they are losing out in terms of business and profitability to the smaller firms. They don't get the trades done anymore because someone else is faster and better in the market. So they then do some research in terms of finding out who that is and what they do better, and try to catch up. Then that usually filters through to the customers of the investment banks, which are the major cutting-edge hedge funds that get the insight on what the banks are doing now. Those hedge funds try to replicate it and it goes to bigger hedge funds, then finally it becomes common knowledge, and people start talking in conferences and it gets dispersed to the general public and the big asset managers. So that's the national progression I would see in technology. (Interview with Peter Van Kleef, High Frequency Trading Review, Sept 2010.)

We are left with incompatible views of, on the one hand, a few superfast players sheltered by barriers to entry and increasing adverse selection for everybody else, against an opposed view of the world where outfits that are small in terms of capital or number of staff dominate the field and make markets more contestable not through financial firepower but their ability to attract and combine some of the best skills available. Note that empirical research such as Brogaard (2010) identifies only a small number of HF firms, but more hard evidence in this area would be important.

4.2. Who is showing the most concern in these debates?

Anecdotal evidence collected in interviews or from the financial press suggests that the institutional buy side is the constituency that feels the most obviously aggrieved by HF activities. This can be illustrated by the responses to a recent regulatory consultation (IOSCO 2011). In the consultation document sent out, questions 9 and 10 were relevant to market abuse:

> Q9: Do you think existing laws and rules on market abuse and disorderly trading cover computer generated orders and are relevant in today's market environment?
>
> Q10: Are there any strategies employed by HFT firms that raise particular concerns? If so, how would you recommend that regulators address them?

Among the entities that responded to Q9, all those that responded 'No', thereby indicating that they felt inadequately protected from new forms of abuse inflicted by HF players, were bodies representing the buy side. Similarly, overall replies to Q10 were split, except among the institutions representing the buy side, where the answer was overhelmingly 'Yes', followed by description of the specific concerns.

We now turn to the link between HFT and market abuse. Most of what follows will be about manipulative strategies and not insider trading. We have not seen the argument made that HF

trading on its own could make insider trading more prevalent. HFT could, of course, make the profits from exploiting a particular piece of inside information larger though.

5. HFT and abuse 1: speed and abusive strategies

It is almost an accepted view amongst many regulators and market commentators that a speed advantage has allowed some firms to conduct new forms of manipulation. Regulatory and academic discussions of HFT produce lists of colourful names describing alleged manipulative HF strategies. There is mention of 'quote stuffing', 'smoking', 'spoofing', 'momentum ignition', 'book layering', etc. The interpretation of some of these terms varies across documents, suggesting that the meaning of this terminology is clearly not fixed. Sometimes 'order pinging' and 'flash quotes' are also mentioned as part of attempts to manipulate.[18]

Several of these tactics do not seem new and, by the same token, nor is it obvious that they require low latency. Those described as 'momentum ignition' or 'book layering' that attempt to lure traders on one side or the other of the market by artificially creating temporary order flow imbalances, or to exploit the predictable reaction of momentum types by exerting price pressure seem like the exploitation of 'positive feedback' traders that was modeled over 20 years ago by DeLong et al. (1990).[19] There are many stories among older London traders of situations where a group of brokers would try to trigger price momentum in times much before HFT. One such story involves traders simultaneously phoning several market makers and hitting them with orders on one side, hoping that this caused a big revision in price and then reversing their trades at favourable prices. Using small orders to detect hidden liquidity (or to find the identity of traders in a non post-trade anonymous market) was surely done for a long time and just not what we now call 'pinging'. Other strategies are just order anticipation, now often called "Predatory Trading" that we know well are not new (see Friederich and Payne 2011b, who supply evidence consistent with predatory trading in pre-HFT markets).

An exception may be quote stuffing, though its incidence is unclear. A controversy was sparked in 2010 by the data firm Nanex when it claimed to have uncovered evidence of quote stuffing on the markets.[20] Not everybody is convinced. Hasbrouck and Saar (2010) note that "our observation that there is no change in the pattern of executions during or immediately after many of these episodes suggests that the story behind this phenomenon may be more complex". Angel (2011, p. 8) believes quote stuffing patterns to be consistent with a "race condition" between algorithms.[21] Misra (2011) feels that these practices "definitely exist" but that they are "blown out of all proportions".[22] He indicates that quote stuffing should be easily

[18] U.S. Senator Charles Schumer has called for the SEC to ban flash quotes, and the European Parliament has denounced them as manipulative.

[19] Cherian and Jarrow (1995, p. 626) note: "Strategies like stop-loss orders, portfolio insurance, technical analysis, etc., are examples of the positive feedback types".

[20] This controversy is ongoing. See <http://www.nanex.net/StrangeDays/08252011.html> or the Interview with Eric Scott Hunsader, "Fighting High-Frequency spam", *The High-Frequency Trading Review,* 26 Oct 2011.

[21] Interview with James J. Angel, "HFT and the Fairness of Markets", *The High-Frequency Trading Review*, 19 July 2011.

[22] Interview with Hirander Misra, "The good and bad of HFT…", *The High-Frequency Trading Review*, 9 September 2011.

detected and dealt with by the trading venues themselves that should have a strong commercial incentive to supply orderly trading services.

Moreover, descriptions of HF players easily taking advantage of others don't always note how risky these strategies can be to those undertake them, because there are other very low latency players out there. Whilst regognising that quote stuffing occurs, Misra points out that it is a crude strategy and those who undertake it are "open to getting hit by more sophisticated high frequency trading firms. All of a sudden they could be sitting on a price movement and massive positions." (Interview mentioned in previous footnote). That these tactics are risky clearly does not mean that they are not manipulative. However, quote stuffing would only create arbitrage profits in a world with one HF player, and this may go to limit its extent.

On the whole, this doesn't seem to us to provide a case to amend existing definitions of abuse at this stage. We are not convinced that most of these practices are very new and that, if documented, they would be classified unambiguously as manipulative using the definitions given early on in this Report. This doesn't preclude them being more prevalent or profitable than they were in the past, but no firm has a monopoly on ultra-low latency.

A separate line of reasoning says that abuse may be easier to conduct in HFT-dominated markets.

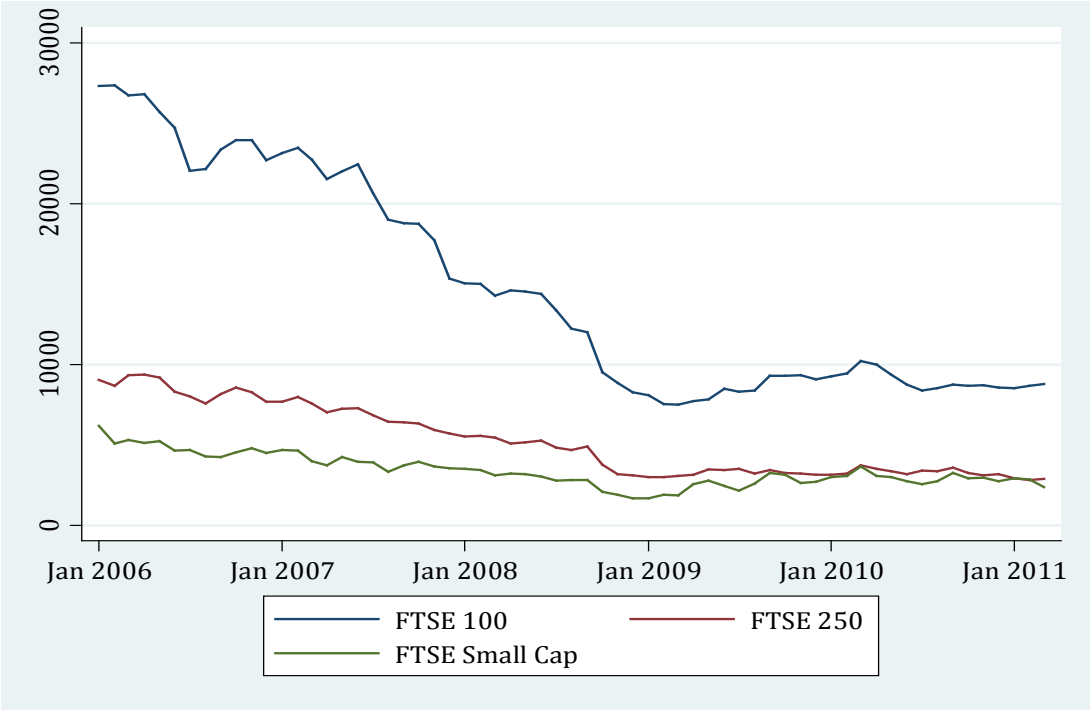
6. HFT and abuse 2: has HFT altered the trading environment in ways that facilitate abusive strategies?

Are old abusive strategies more prevalent, more profitable or harder to catch than before in today's markets? In this section, we discuss changes in trading that may be attributable to HFT and be associated with a higher incidence of abuse or perception thereof, whilst *not* implying that HF traders take advantage of slower players.

6.1. The new order flow

HFT may have changed the time-series properties of order flow in ways that we don't yet understand. A dramatic decrease in trade sizes is well documented (see Chordia, Roll and Subrahmanyam, 2011 for the US market or Figure 2, drawn from Friederich and Payne, 2011a, for the UK). This decline could be driven by changes in either or both of the liquidity demand and supply sizes: on the demand side, order execution algorithms are designed to minimise market impact by using worked, dynamic execution. On the supply side, market-marking algos are designed to avoid suffering losses from 'toxic' order flow, whether because it's better informed or just based on 'order anticipation'. This may cause a reduction of their quoted sizes.

Figure 2: The figure shows the average trade size (GBP) on the London Stock Exchange's order book for large, mid and small caps, 2006-2011.



(Source: Friederich and Payne, 2011a)

The role of latency in all this is not clear but Brogaard (2010) finds that HF traders contribute much less depth than do non-HFT agents.

Note that Industrial Organisation aspects may be germane to this debate and they are often overlooked. 'Pure' HFT firms, as distinct from Hedge funds or sell-side firms that implement HF strategies are a new type of entity in a market structure (not microstructure) sense. The way they trade is an integral part of their business and financial model:

- Because the positions they take are small and short-lived, they can set up shop with a comparatively modest amount of capital, without needing to raise additional capital from outside investors. This should create competition for the large, established sell side and contribute to making markets more contestable.

- Because they are not open to outside investors they can also remain pure proprietary traders.

- However, their limited capital is perhaps among the reasons why they are unable to supply vast amounts of depth.

The shallow order flow generated by HF traders may make the work of manipulators easier. If HFT equals lower spreads but also lower depth or resilience, then it is possible that abuse becomes easier as price moves are easier to engineer. Moreover if in risky times, HFTs are induced to unwind positions, due to their capital constraints, this may reinforce price moves and further destabilize markets.

Another feature of the 'new' order flow may be its degree of autocorrelation in direction – the

result of algos breaking up and 'working' large orders electronically. We come back to this on the next page.

Finally, order flow fragmentation, which spreads liquidity across several venues, and the decrease in tick sizes, which decrease quoted depth, are potential aggravating factors. We come back to these points below.

6.2. Price pressure effects and the "Predatory Trading" debates

A specific example of how shallow order flow could foster opportunistic behaviour is *predatory trading,* sometimes also described as "order anticipation" or "liquidity detection". *The Trade News* (26 August 2011) reports that "Some 25% of poll respondents [on the buy side] considered that predatory high-frequency trading (HFT) algorithms represent the most severe risk of information leakage" (although this is a distant second to "Sales trader indiscretion").

By "predatory trading", we refer specifically to the strategy modeled in the work of Brunnermeier and Pedersen (2005). In that sense, predation is a strategy of trading at the same time and with same direction as a large investor whose presence has been detected by their 'footprints' in orders and trades. Concomitant trading by the predator amplifies the price pressure exerted by the investor's own trades, and the strategy is settled at a profit by the predator by reversing their trades when they sense that the large investor has left the market.

Friederich and Payne (2011b) find that the shallower and the more autocorrelated in direction the order flow is, the more profitable predation can be because price pressure is more easily exerted and forthcoming order flow is more predictable. This ties in with the alleged characteristics of HFT-dominated order flow.

Whether predatory behaviour represents manipulation could be open to debate:

- In theory, Brunnermeier and Pedersen (2005) state that "The notion of predatory trading partially overlaps with that of stock price manipulation" (p. 1828) but it is distinct from it because strictly speaking, predation cannot be conducted by the predator only, as it exploit a large trader's price pressure. The model, however, does not consider the possibility that predation can be conducted by liquidity suppliers though quote "shading" or "fading" (Angel et al. 2011), which as a deliberate attempt to create a temporary price change to profit from, seems more akin to outright manipulation. The argument relies on concentration in liquidity supply, again an alleged characteristic of HFT-dominated order flow. In the related paper by Attari et al. 2005, the authors explicitly describe the strategy as "manipulation".

- Empirically, Friederich and Payne (2011b) document that predatory trading may damage market liquidity if the conditions giving rise to predation occur frequently (as opposed to the extreme situations of financial distress envisaged by Brunnermeier and Pedersen). Thus predation has some of the market-wide implications of outright manipulation.

- In regulatory terms, under the definitions given above, it may constitute manipulation (as a type of behaviour that "gives a false or misleading impression of either the supply of, or demand for, an investment" etc.).

    - The SEC showed concern by recently calling for evidence on order anticipation strategies, described as "any means to ascertain the existence of a large buyer (seller) that does not

involve violation of a duty (...) or other misconduct." The SEC further notes that "Order anticipation is not a new strategy. (...) Do commenters believe that order anticipation significantly detracts from market quality and harms institutional investors (...)?" (SEC 2010, pp. 54-56.)

- And IOSCO (2011) similarly notes that a "(…) specific concern is whether HFT enables the pursuit of some apparently non-abusive practices, such as searching for hidden liquidity and order anticipation, that results in harm to market quality and prejudices confidence in markets if undertaken on a large scale."

The fear of predation, as distinct from its actual incidence, could be enough to affect agents' willingness to expose their orders to the market and cause a deterioration of liquidity. If considered distortionary or manipulative behaviour along the Kyle-Viswanathan definition, we agree that it would be a tall order for regulators to identify predation conclusively, but perhaps not a taller order than most other forms of abuse.

6.3.A poorer "market ecology"?

Speed may have caused a reduction in the number and diversity of liquidity suppliers and their strategies. As a result of HFTs more or less always beating the others to the post in terms of order submission, other agents are reduced to consuming liquidity. Liquidity supply becomes dominated by a few firms with a similar agenda. Anecdotal evidence is consistent with this view. A broker describes HF traders as "crowding out the sell-side and the institutions in terms of being able to earn spreads and rebates [through liquidity supply]."[23] Another participant finds that "There has been a loss of 'diversity in opinion' which is a desired characteristic for a market."[24] Hunsader concurs: "When everything is about speed, you lose a lot of diversity; you don't have all those different players with different viewpoints using different algorithms and with different strategies at play. You end up killing all of them (…)."[25]

Exchange fee structures may play a part in this. Volume rebates and maker-taker fees influence the strategies of HFT players by encouraging them to attempt to earn the spread, perhaps elbowing out other liquidity suppliers in the process.

There is a scenario where markets are increasingly dependent on a handful of the lowest latency firms that dominate liquidity supply by dint of sheer speed. Being able to revise orders faster, they leave no choice to other traders but to be liquidity consumers – they are effectively excluded from liquidity supply. This separates agents demanding and supplying liquidity more strictly than before and may make for an order book where liquidity is more fickle. Moreover, average HFT depth may be limited for reasons described above and these firms may step away from the top of the order book at the slightest hint of unusual order flow ('toxic' flow or attempted manipulation). This may give free rein to manipulators to exert price pressure much more easily than was the case before.

---

[23] Interview with Joe Gawronski, "The Changing Role of the Broker", *The High-Frequency Trading Review*, 28 September 2010.

[24] Interview with Karim Taleb, "How High Frequency Trading is Changing the Markets", *The High Frequency Trading Review*, 10 September 2010.

[25] Interview with Eric Scott Hunsader, "Fighting High-Frequency spam", *The High-Frequency Trading Review,* 26 Oct 2011.

6.4. Automation, the reduction of the strategy space, and predictability

Another source of abuse may lie in the fact that trading processes are more automated than they've ever been. At the same time, some HFT algorithms are not terribly complex and may react in predictable ways to certain market conditions. Hence, although there is much variety to the strategies that are conducted via computer, there is predictability in the behaviour of at least some algo scripts that can be exploited.

O'Hara (2010) gives an illustration that involves two market-making scripts where one games the other, driving prices away from equilibrium. The CEO of a firm that sells execution software to the buy side notes that "If you think about how execution algorithms work, some behave in such a formulaic way that even human traders can sit there watching a real-time price chart and see exactly when the algo kicks in (…). Now if a human being can see it, you can be sure that a machine can spot it too (…)."[26] The focus on "execution algorithms" here may point to incentives issues -- there may be more incentives to develop the most efficient and aggressive scripts to conduct prop trading than for client execution. In its contribution to IOSCO's consultation in 2011, Deutsche Boerse indicates that "Concerning the pattern "layering" we would like to note that this behaviour is in general not conducted by machines, but by pure manual trading or partially "electronically assisted" manual trading. Implemented algorithms are easily ignited by a layered order book as such delinquents enter non bona fide orders to mislead the market and in particular a specific implementation of an algorithmic trading strategy. The delinquents do take in general advantage of inappropriate programmed algorithm."

6.5. Has HFT rendered market abuse much harder to detect?

Increase in traffic The daily terabytes of quotes and trades data generated by HFT may be overwhelming. ESMA (2011) states: "When thousands of order messages a second are flowing to individual trading platforms it increases the challenge of spotting potentially abusive behaviour. Efforts to compress data to make it more manageable through aggregation are likely to hide much more than they reveal and the number of false positives produced by systems designed to provide alerts will rise."

Increased demands on regulators in trading and IT sophistication Keeping up with the technical aspects of trading is a challenge to all regulators. A report by the Boston Consulting Group recently warned that the SEC had to substantially improve its IT and trading sophistication to stand any chance of detecting market abuse.[27] The report notes: "The agency does not have sufficient in-house expertise to investigate the inner workings of [trading] algorithms (…)" and recommends that "The SEC should have staff who understand how to (…) perform the analytics required to support investigations and assessments related to high-frequency trading. (…) the SEC requires people who know how HF traders use technology."

Both reasons may rationally cause a decline in investors' confidence in the fairness of markets. It was perhaps an extreme case (and not one of abuse) but it took American regulators five months to analyse 20 minutes of data covering the Flash Crash of May 2010, and we have still

[26] Interview with Marcus Hooper, "Can Predictive Algorithmic Switching be the Buy-Side's Answer to High Frequency Trading?", *The High Frequency Trading Review*, 9 May 2011

[27] "U.S. Securities and Exchange Commission, Organizational Study and Reform", Boston Consulting Group report, March 10, 2011.

not come to an agreed view of what happened. This should be a worrying illustration of the difficulty that exists in conducting investigations involving today's data.

6.6. Informational aspects

An HFT-driven *decrease in pre-trade transparency* may also give investors the feeling that they are put upon. Large investors complain about the difficulty of knowing what the current price is when quotes are flashed faster that the human eye can see or they continuously 'flicker'. The institutional buy side in particular voices frustration with the "wall of HFT noise" which would leave them at the mercy of their brokers and HF liquidity suppliers, who know "where the market is" much better than they do.

Different participants will not only see a different picture of total order flow, they will also receive it as different speeds depending on how their IT infrastructure is designed, how physically close they are to the exchanges or data vendors, etc. In a world where milliseconds matter, there is likely to be truth to the claim that different agents see a different market to a greater extent that was the case before. A Director of the MTF Turquoise puts it this way:

> Where a customer is, relative to the different market centres, means that their view of price (and more importantly, not just their view of what the current price is, but actually their ability to trade against those prices) is totally different to a customer located somewhere else. I think we have to abandon this idea that there is a universal truth for the best currently available price. In some ways the consolidated quote in the United States is actually no longer real. The reason no professional market participant uses it for trading and submission of orders is because it's wrong whoever you are. (Interview with Natan Tiefenbrun, The High Frequency Trading Review, 3 December 2010.)

Whilst an increase in the difference of the perception of the market across agents may be disturbing to some, it is nothing new and it can't represent abuse. Traders closer to the real prices (either physically or temporally) have always attempted to take advantage of those further away, and the speed of accessing and trading on market data was always a function of the investments that a firm has undertaken in that respect.

7. HFT and abuse 3: have confounding factors increased the *perceived* likelihood of abuse?

We now discuss changes in the trading environment that may have affected the perception of abuse in today's markets while being less clearly or not at all attributable to HFT growth.

7.1. The great tangle: latency and confounding factors

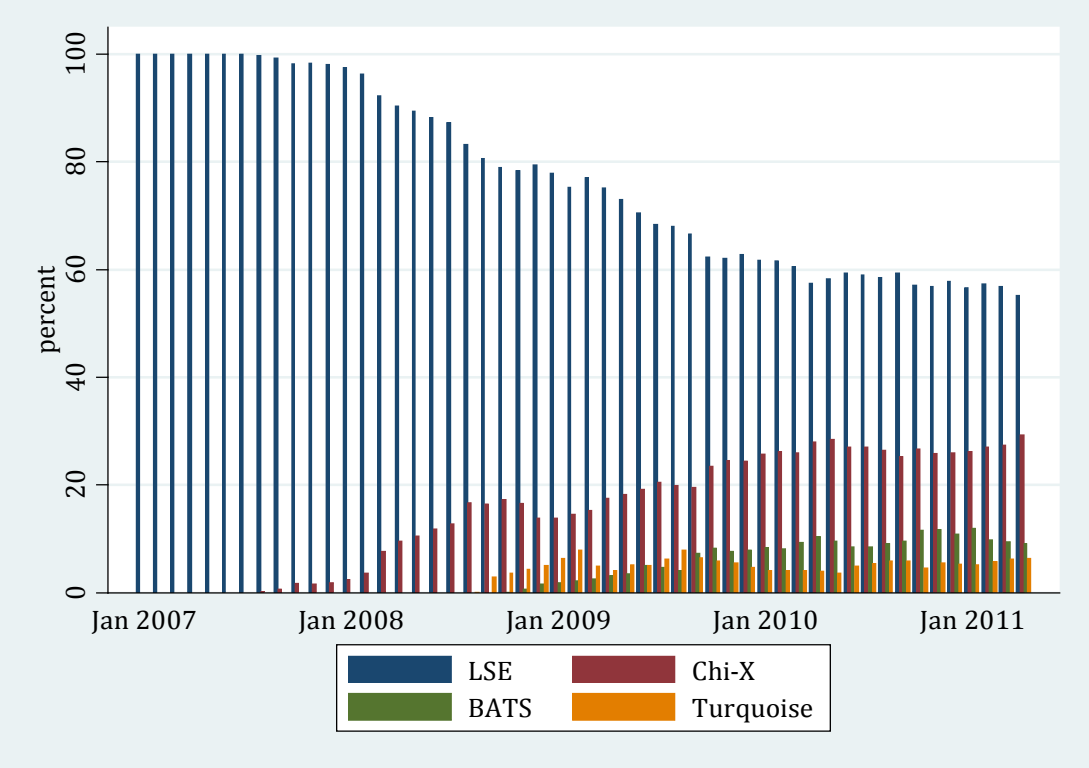
A previous section outlined the fact that discussions of HFT should attempt to isolate the role of low latency from that of deep *preceding* trends such as the end of formal market-making commitments, the automation of trading processes and the growth in proprietary trading (even though it may have accelerated some of these trends). To continue on a related thread, it is also the case that some developments that were *contemporaneous* to the growth in HFT may make the effect of latency as such hard to identify.

7.2.Fragmentation

Fragmentation of *order flow* deserves special mention in this respect. That is, many venues and modes of trades are now available to traders, especially institutional ones. Figure 3 below shows the recent growth in order flow fragmentation for the London equity market:

LSE, CHI-X, BATS AND TURQUOISE MARKET SHARES BY VALUE TRADED,
FTSE 100 STOCKS, 2007-APRIL 2011

Figure 3: the recent growth in order flow fragmentation for the London equity market.



(Source: Friederich and Payne, 2011a)

Therefore the increase in fragmentation was to a fair extent contemporaneous with the growth of HFT, although fragmentation started later (round early 2008) and seems to have stabilized in late 2009, whereas the estimates of HFT we have indicate continuous growth over the period 2007-2010.

Compare the current trading landscape to that of a few years ago: the London Stock Exchange had a near monopoly on UK stock trading, and the top ten (human) brokers would intermediate perhaps 80% of all order flow on the order book at much slower speed. Today's trading environment is more complex in every way: a larger number of participants intervene on a dozen venues instead of one, all this at speeds that make these activities seem simultaneous in time to humans but not at all to agents that operate in milliseconds.

The difficulty to understand today's trading structure may add to the frustration and the sense of alienation from the trading process felt by some investors. Fragmentation of order flow within and across borders is distinct from HF trading but it certainly makes the job of regulators harder and may encourage unscrupulous traders.

7.3. Informational aspects

First, order flow fragmentation generates fragmentation in an *informational* sense in European markets, as there is no consolidated tape displaying prices across all trading venues as in the US. Different market participants must decide from which vendor they will purchase their order flow feed and how it is going to be displayed by choosing a software front end. It required significant investment and is another source of frustration for participants on the buy side below the largest funds, who find it ever harder to assess the quality of execution they receive from their broker pre and post trade. It may add to the sense of "uneven playing field".

A sense of *reduced pre-trade transparency* may also be attributable to the growth in new types of orders (hidden or 'iceberg') specifically designed to hide a substantial part of one's trading interest, and the availability of 'Dark' venues, also designed to avoid revealing one's intentions, was largely concomitant with the growth in HFT.

These developments may have been accelerated by HFT but they are clearly distinct from it and they have preceded it to a fair extent.

Reduction in tick sizes: this is another development that's tied to HFT growth but is distinct from it. Smaller tick sizes influence order submission behaviour and are generally considered to have a negative impact on quoted depth. Lower average depth is often attributed to HF strategies but may also be due to reduced tick sizes.

Discussing alleged momentum-based manipulation, the Australian Stock Exchange (ASX 2010) recognises that it may be less to do with HFT than fragmentation by cautiously stating: "(…) some so-called 'momentum' algorithms have the potential to distort price discovery for a security. Within a multi-market operator environment, where liquidity has been fragmented and where maker-taker pricing encourages algorithms to 'chase' one another to receive incentives, the risk of price distortion increases significantly."[28]

8. Can HFT make abuse *less* likely?

This view doesn't seem to be often entertained. Yet the case can be made that at least some HFT activity has the potential to make manipulation attempts more risky and costly, and therefore to reduce their incidence.

As described above, a sizeable subset of HF traders are statistical arbitrageurs whose focus is the relative pricing of economically similar securities. As such they should trade so as to contribute to keeping the prices of these similar assets in equilibrium and thus provide a counter-balancing force to those attempting to manipulate. The manipulator, for example, may

[28] Australian Stock Exchange, "Review of Algorithmic Trading and Market Access Arrangements", 8 February 2010.

try to move a stock price downwards through a burst of aggressive selling. A stat-arb algorithm would immediately spot that this stock's price has moved out of line compared to its peers' prices, and would quickly buy the manipulated stock (while selling some of the peer group). This dampens the price move and makes it harder for the manipulator to trade at a profit. If the manipulator was able to affect the prices of the entire set of stocks simultaneously then the stat-arb model could not correct the mis-pricing, but this would require the commitment of significant capital and a sophisticated manipulative algorithm.

Thus, HFT algorithms of the stat-arb and market-making types are likely to limit the scope for manipulation. It is not clear which dominates, on average, between the forces of equilibrium and manipulation.

9. How does market surveillance adapt to an HFT world?

Aitken notes that in terms of sophistication and therefore ability to conduct illegitimate trading, "Regulators are behind exchanges and exchanges are behind brokers."[29] Market surveillance is gradually adapting to changes in trading practices though, whether they are HFT-driven or not.

- At the individual firm level, software is developed that assesses the toxicity of order flow on each venue by examining patterns in recent activity. A green light is given when it's safe.[30]

- At the level of trading venues and brokers, there is increased reliance on other traders to report potential abuse and on trading venues to police it. Some years ago, the historical exchanges held a near monopoly on trading in their shares and could perhaps afford to be more complacent. Today, the operators of electronic trading venues have stronger incentives to deal with simple forms of abuse because they are competing to attract order flow and must be seen to maintain an orderly trading system. Other traders should have equally strong incentives to report wrongdoers to the venues.

- At the level of national regulators, recent steps taken also rely on traders. In Europe, this is formalised in the "Suspicious Transactions Reports" (STRs) stemming from the Market Abuse Directive and implemented in the UK since 2005.

- Across markets: surveillance across venues is made crucial by the extent of fragmentation and the fact that abuse can be conducted across borders, classes of assets. This is one of the thorniest issues facing regulators.

The broad trends are therefore to place more emphasis on market participants themselves ('micro-surveillance'), to improve regulator ability to track traders across markets through the use of unique trader identifiers, and to try and implement a degree of real-time surveillance to weed out abuse as early as possible. The data overload problem noted above is here to stay though. ASX (2010) reports that "Several market supervisors indicated that the split between real time and post trade analysis is roughly 5%:95%. This strong emphasis on post-trade

[29] http://video.ft.com/v/89326741001/Jun-1-Can-regulators-keep-up-

[30] See, e.g., "Manipulative trading detection systems spring up", *Wall Street Letter*, April 18, 2011.

surveillance activity is necessary due to the large volumes of data that need to be integrated before sophisticated pattern analysis software can scan the data for indications of misconduct."

In May 2011, US Senator Charles Schumer has suggested that HF traders shoulder the extra cost of market surveillance that their activities generate, although this idea may not be easy to implement.

An ongoing discussion involves the existence of a socially wasteful level of investment associated to HFT (the "arms race" argument). Increased market surveillance cost inflicted on society could even more directly result from HFT.

10. Concluding comments

Like many reports focussing on HFT, the current one contributes at this stage many more questions than answers.

We have attempted to draw attention to two points: first, anecdotal evidence from the financial press, discussion with traders, or reactions to regulatory calls for evidence suggests to us that the perception of abuse or "unfairness" may have increased in HFT-driven markets. The extent to which this perception is at variance with current reality is unclear for lack of evidence on the extent of abuse.

Second, we suggest that this shift in perceptions (and perhaps reality) could have occurred for three sets of reasons:

- HF traders may be able to exploit a latency advantage to inflict losses on slower players. It is unclear to what extent these strategies are sustainable to those who implement them unless they hold a speed advantage over nearly everybody else.

- The arrival of HFT may have altered the trading environment in ways that have increased abuse or its perception: thinner and more correlated order flow, increased opacity, use of more generic strategies prone to manipulation, concentration in liquidity supply. Strategies that benefit from the creation of price pressure effects or from the predictability of algo behaviour may be more tempting to conduct than they used to be. They may also be harder to catch: the technical and data aspects of the new trading environment make market monitoring harder.

- Concomitant changes that were only partly or not at all driven by HFT may have increased the perception of abuse: fragmentation in an order flow or informational sense and reduced pre-trade transparency are examples.

To use an analogy, HF trading is described as a mode of trade that relies on very small average profits that become significant when accumulated over a very large number of trades. Manipulative practices may have followed suit: the possibility that 'micro-manipulation' relying on the high-frequency generation of small, temporary price deviations may be more commonplace than before ought to be entertained.

A countervailing force is that HFT arbitrageurs have powerful means and incentives to keep

prices in equilibrium. This argument is often disregarded.

Beyond a handful of publicized cases that were prosecuted by regulators, there is no evidence on the reality of abuse in HFT-dominated markets. Industry voices on all sides are calling for more to be produced. "*If the regulator has concerns over possible breaches of the market abuse regime in Europe as it currently stands, research into this area should be initiated. Although time consuming and costly, it may help to put to rest wider concerns if conducted effectively.*" (AFME 2011)

References

AFME (2011), Response to IOSCO Consultation Report on 'Regulatory Issues Raised by the Impact of Technological Changes on Market Integrity and Efficiency', Association for Financial Markets in Europe, 12 August 2011.

Angel, James J. and Douglas McCabe (2010), "Fairness in Financial Markets: The Case of High Frequency Trading", McDonough School of Business Working Paper, December.

Angel, James, Lawrence Harris and Chester Spatt (2011), "Equity Trading in the 21st Century," *Quarterly Journal of Finance*, 1(2), 1-53.

ASX (2010), "Review of Algorithmic Trading and Market Access Arrangements", Australian Stock Exchange, 8 February 2010.

Attari, Mukarram, Antonio S. Mello, and Martin E. Ruckes (2005), "Arbitraging Arbitrageurs", *Journal of Finance*, 60(5), 2471-2511.

Biais, Bruno, and Paul Woolley (2011), "High-Frequency Trading", Working Paper, March.

Brogaard, Jonathan (2010), "High Frequency Trading and its Impact on Market Quality", working paper.

Brunnermeier, M. K., and L. H. Pedersen (2005), "Predatory Trading", *Journal of Finance*, 60, 1825-1863.

Cartea, Alvaro and Jose Penalva (2011) "Where is the Value In High-Frequency Trading," working paper.

Cherian, Joseph A. and Robert A. Jarrow (1995), "Market Manipulation" in R. Jarrow et al., Eds., *Handbooks in Operations Research & Management Science*, Vol. 9, Ch. 20, Elsevier.

Chordia, Tarun, Richard Roll and Avanidhar Subrahmanyam (2011), "Recent trends in trading activity and market quality", *Journal of Financial Economics*, (DOI: 10.1016/j.jfineco.2011.03.008).

Cohen, Samuel and Lukasz Szpruch (2011) " A Limit Order Book for Latency Arbitrage," working paper.

DeLong, Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann (1990), "Positive-Feedback Investment Strategies and Destabilizing Rational Speculation", *Journal of Finance*.

ESMA (2011), "Guidelines on systems and controls in a highly automated trading environment for trading platforms, investment firms and competent authorities," Consultation paper, European Securities and Markets Authority, 113 pages, 20 July 2011, http://www.esma.europa.eu/popup2.php?id=7675

European Commission (2010), Consultation on the review of the Markets in Financial Instruments Directive (MiFID), 8 December 2010, 83 pages, http://ec.europa.eu/internal_market/consultations/docs/2010/mifid/consultation_paper_en.pdf

Foucault, Thierry (1999) "Order Flow Composition and Trading Costs in a Dynamic Limit Order Market," *Journal of Financial Markets*, 2, 99-134.

Friederich, Sylvain and Richard Payne (2011a), "Computer-Based Trading, Liquidity and Trading Costs", report to Foresight, July, http://www.bis.gov.uk/assets/bispartners/foresight/docs/computer-trading/11-1240-dr5-computer-based-trading-liquidity-and-trading-costs.pdf.

Friederich, Sylvain and Richard Payne (2011b), "Trading Anonymity and Predatory Behaviour", Working Paper, September.

Gerig, Austin and David Michayluk (2010) "Automated Liquidity Provision and the Demise of Traditional Market Making,", working paper.

Glosten, Lawrence and Paul Milgrom (1985) "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders", *Journal of Financial Economics*, 14, 71-100.

Harris, Lawrence and Ethan Namvar (2011), "The Economics of Flash Orders and Trading", Working Paper.

Hasbrouck, Joel and Gideon Saar (2010), "Low-Latency Trading", Johnson School Research Paper, October.

Hoffman, Peter (2011) "Algorithmic Trading in a Dynamic Limit Order Market", working paper.

IOSCO (2011), "Regulatory Issues Raised by the Impact of Technological Changes on Market Integrity and Efficiency", International Organization of Securities Commissions, October 2011, http://www.iosco.org/library/pubdocs/pdf/IOSCOPD361.pdf.

Jarrow, Robert and Philip Protter (2011) "A Dysfunctional Role of High-Frequency Trading in Electronic Markets," working paper.

Jovanovic, Boyan and Albert Menkveld (2010) "Middlemen in Limit Order Markets", working paper.

Kyle, Albert S. and S. Viswanathan (2008), "How to Define Illegal Price Manipulation", *American Economic Review*: Papers & Proceedings, 98:2, pp. 274–279.

Martinez, Victor and Ioanid Rosu (2011) "High Frequency Traders, News and Volatility," working paper.

O'Hara, Maureen (2010), "What is a Quote?", *Journal of Trading*, Spring.

Putnins, Talis J. (2011), "Market Manipulation: a Survey", *Journal of Economic Surveys* doi: <10.1111/j.1467-6419.2011.00692.x>

SEC (2010), “Concept Release on Equity Market Structure”, US Securities and Exchange Commission, 14 January 2010, mimeo, 74 pages, http://www.sec.gov/rules/concept/2010/34-61358.pdf

Zhang, Frank and Stuart Baden Powell, “The Impact of High-Frequency Trading on Markets”, *CFA Magazine*, March/April 2011.

Foresight
1 Victoria Street
London SW1H 0ET
www.foresight.gov.uk

URN: 12/1053