Government Office for Science

Foresight

The Future of Computer Trading in Financial Markets

Working paper

Foresight, Government Office for Science

This working paper has been commissioned as part of the UK Government's Foresight Project on The Future of Computer Trading in Financial Markets. The views expressed are not those of the UK Government and do not represent its policies.

Introduction by Professor Sir John Beddington

Computer based trading has transformed how our financial markets operate. The volume of financial products traded through computer automated trading taking place at high speed and with little human involvement has increased dramatically in the past few years. For example, today, over one third of United Kingdom equity trading volume is generated through high frequency automated computer trading while in the US this figure is closer to three-quarters.

Whilst the prevalence of computer based trading is not disputed, there are diverse views on the risks and benefits which it brings today, and how these could develop in the future. Gaining a better understanding of these issues is critical as they affect the health of the financial services sector and the wider economies this serves. The increasingly rapid changes in financial markets mean that foresight is vital if a resilient regulatory framework is to be put in place. A key aim of this Foresight project, which has been overseen by a group of leading experts, has therefore been to draw upon the very best science and evidence from across the world to take an independent look at these issues.

The three papers presented here review evidence directly commissioned by the project as well as the wider evidence base. Leading experts from over 20 countries have been involved in writing and peer reviewing this material. The first paper considers the effect of computer trading on financial stability. It reviews the evidence of its past effect, and considers possible future risks. In contrast, the second paper considers the benefits that computer trading has had on liquidity, price efficiency and transaction costs. Together these two papers paint a picture of both risks and benefits, and for this reason neither paper should be read in isolation. The third paper focuses on technology.

Importantly, the results documented here represent the independent views of academics. In particular, this working paper does not represent the position of the UK or any other government, nor does it seek to further the interests of any part of the financial services sector. However, whilst these papers are not presented as the last word on the topics they address, I hope that they can make a substantial contribution to current debate. It is on this basis that I take great pleasure in making them freely available.

Professor Sir John Beddington CMG, FRS
Chief Scientific Adviser to HM Government and
Head of the Government Office for Science

Lead Expert group overseeing the project

Dame Clara Furse DBE – Non-Executive Director, Nomura

Andy Haldane – Executive Director for Financial Stability at the Bank of England

Professor Charles Goodhart – Professor of Finance, London School of Economics

Professor Dave Cliff – Professor of Computer Science, University of Bristol

Dr Jean-Pierre Zigrand – Reader in Finance, London School of Economics

Kevin Houstoun – Chairman, Rapid Addition; Co-Chair, Global Technical Committee, FIX Protocol Limited

Professor Oliver Linton – Chair of Political Economy, Cambridge University

Professor Philip Bond – Oxford Centre for Industrial and Applied Mathematics

Foresight project team

Sandy Thomas (Head of Foresight)

Derek Flynn (Deputy Head of Foresight)

Lucas Pedace (Project Leader)

Luke Ryder (Project Manager)

Chris Griffin (Project Manager)

Piers Davenport (Executive Officer)

Yasmin Hossain (Research Officer)

Isabel Hacche (Intern)

For further information about the project please visit:

http://www.bis.gov.uk/foresight/our-work/projects/current-projects/computer-trading

Contents

The Future of Computer Trading in Financial Markets 1

Introduction by Professor Sir John Beddington 2

Lead Expert group overseeing the project:..3

Foresight project team:..3

For further information about the project please visit: 3

Paper 1: Financial stability and computer based trading 6

Key findings 7

Executive summary 8

1. Introduction...9

2. How has CBT affected financial stability in the past?..9

3. How is CBT thought to affect financial stability?...13

3.1 Mechanisms of instability..13

3.2 Interactions between mechanisms 17

3.3 Socio-technical factors: normalisation of deviance...19

4. How can CBT affect financial stability in future? 21

5. Conclusion..23

Paper 2: The impact of computer trading on liquidity, price efficiency/discovery and transaction costs 24

Key findings 25

Executive summary 26

1. Introduction...27

2. The impact of computer trading on liquidity, transaction costs, and price efficiency...27

3. Past: what has been the impact of computer trading on market quality in recent years?...30

3.1 Liquidity 30

3.2 Transaction costs 33

3.3 Price efficiency 34

4. Present: what has been the impact of computer trading on market quality at present?..35

5. Future: how is the impact of computer trading on liquidity likely to evolve in the next 10 years? 37

6. Conclusions 38

Paper 3: The impact of technology developments 40

Key findings 41

Executive Summary...42

1: Introduction...43

2: How has financial market technology evolved?...43

3: What are the key current technology developments? 46

3.1 Automated analysis of market news and sentiment 48

3.2 Studying interactions between human and algorithmic trading systems 49

3.3 From the present to the future 50

4: Technology advances likely in the next 10 years 51

4.1 Cloud computing...51

4.2 Custom silicon 52

4.3 Computer-generated trading algorithms that adapt and learn 53

5: Conclusion...53

Glossary of terms 54

Paper 1: Financial stability and computer based trading

Jean-Pierre Zigrand, London School of Economics,

Dave Cliff, University of Bristol,

Terrence Hendershott, University of California at Berkeley.

Key findings

Economic research thus far provides no direct evidence that high frequency computer based trading has increased volatility.

However, in specific circumstances, a key type of mechanism can lead to significant instability in financial markets with computer based trading (CBT): self-reinforcing feedback loops (the effect of a small change looping back on itself and triggering a bigger change, which again loops back and so on) within well-intentioned management and control processes can amplify internal risks and lead to undesired interactions and outcomes.

The feedback loops can involve risk-management systems, and can be driven by changes in market volume or volatility, by market news, and by delays in distributing reference data.

A second cause of instability is social: a process known as *normalisation of deviance*, where unexpected and risky events come to be seen as ever more normal (e.g. extremely rapid crashes), until a disaster occurs.

Executive summary

This paper reviews the findings of the “driver review” documents commissioned by the Foresight project, and other recent literature, that explore the effects of computer based trading (CBT) on the stability of financial markets both now and in the future.

We concentrate here on stability as a primary factor affecting confidence in capital markets, including their role as a wealth store or as a borrowing vehicle. Changes and fluctuations in market values are always to be expected, but if a change is sufficiently large or unexpected that it fundamentally impairs the saving/investment process, eroding confidence, then that change can be considered as a financial stability event.

For example, despite being an intra-day event, the “Flash Crash” of 6 May 2010 (when the US equity market dropped by 600 points in 5 minutes, eliminating approximately US$800bn of value, and then regained almost all of the losses within 30 minutes) helped to erode confidence in stock markets sufficiently to be followed by several months of outflows from retail mutual funds in the US.

We identify three main mechanisms that may lead to instability when financial markets involve significant proportions of CBT: nonlinear sensitivities to change (where small changes can have very big effects), incomplete information (where some agents in the market have more, or more accurate, knowledge than others), and internal “endogenous” risks based on feedback loops within the system.

The internal chains of cause and effect that are central to endogenous risks in CBT-markets involve positive feedback loops that can amplify and self-reinforce detrimental interactions between different aspects of well-intentioned management and control processes.

The feedback loops can even be worsened by risk-management systems, and can be driven by changes in market volume or volatility, by market news, and by delays in distributing reference data.

A fourth cause of instability is social: a process known as *normalisation of deviance*, where unexpected and risky events come to be seen as ever more normal, until a disaster occurs.

Finally, in a world with multiple trading and pricing venues that are interconnected by CBT systems, the network topology determines the stability and the flow of information and trades, and hence is a major factor in determining overall systemic stability.

The novel aspects of the dynamics of markets with significant proportions of computer based high frequency traders include: (a) that interactions are taking place at a pace where human intervention could not prevent them –

an important speed limit has been breached; (b) that, given this, computer based (and therefore mechanical) trading is almost obligatory, with all of the system-wide uncertainties that this gives rise to; (c) that information asymmetries then become more acute (and indeed different in nature) than in the past; and (d) that the source of liquidity provision has changed, to computer based and high-frequency trading, which has implications for its robustness under stress.

1. Introduction

We take a broad interpretation of Computer Based Trading (CBT). A useful taxonomy of CBT was proposed in DR5[1], which identifies three characteristics that can be used to classify CBT systems. First, CBT systems can trade on an agency basis (i.e. attempting to get the best possible execution of trades on behalf of clients) or a proprietary basis (i.e. trading using one's own capital); second, CBT systems may adopt liquidity-consuming (aggressive) or liquidity-supplying (passive) trading styles; and third, they may be classified as engaging in either uninformed or informed trading. Much of the current public debate is concerned with the class of aggressive predatory algorithms, especially those that operate at high speed and with high frequency. Because most financial institutions that operate CBT cannot be neatly assigned to only one of the above buckets, it is more fruitful to think about CBT systems, the algorithms they employ directly, and the frequency at which they trade, rather than to think about the behaviour of a particular named financial or trading corporation such as a specific investment bank or fund-management company. For much the same reasons, in the discussion that follows we will not focus on any one asset class (such as equities, foreign exchange, commodities, or government bonds) in particular, but rather we will spell out the forces that seem likely to shape the future stability issues arising from CBT. In this report we summarise the intuition behind some of the more economically plausible risk factors of CBT: these "risk drivers" can best be viewed as forming the logical basis of possible future scenarios concerning the stability of the financial markets.

2. How has CBT affected financial stability in the past?

The *raison d'être* for financial markets is to aggregate myriad individual decisions and to facilitate an efficient allocation of resources in both primary and secondary markets[2] by enabling timely and reliable reaping of mutual gains from trade, as well as by allowing investors to diversify their holdings.

---

[1] Throughout this document DR refers to driver review studies commissioned by the lead expert group. These can be found in the project's webpage: http://bis.gov.uk/foresight/our-work/projects/current-projects/computer-trading

[2] When a company issues equities (shares) to raise capital, that is the primary market in action. When the shares are then subsequently traded among investors and speculators, that is the secondary market in action.

As with many other aspects of modern life, innovations in technology and in finance allow the repetitive and numerically intensive tasks to be increasingly automated and delegated to computers. Automation, and the resulting gains in efficiency and time, can lead to benefits but also to social costs. This paper focuses solely on possible repercussions of CBT on financial stability, especially the risks of instability. This should certainly not be construed as meaning that CBT is socially detrimental or bears only downside risks and costs. The hope is that by better understanding the risk-drivers of CBT on financial stability the creators, users and regulators of CBT systems may be able to manage the risks and allow the benefits of CBT to emerge while reducing the social costs.

The findings in this paper may apply to any given market structure, but we feel they are especially relevant to the continuous auctions of the sort that run on traders' screens in most of the major financial markets worldwide. The reason is that even if daily volume is large, the second-by-second volume may not be. For instance, even a daily turnover of more than $4trn in the foreign exchange market on average corresponds to only $2.7m second-by-second volume for major currency pairs like Euro-USDollar. Even in such a huge market, a sufficiently large order can temporarily sway prices, depending on how many other orders are in the market (the "depth" of the market) at that moment in time.

Price volatility is a fundamental measure useful in characterising financial stability (wildly volatile prices are an indicator of instabilities in the market) [3]. In DR1, Linton and Atak note that, since the turmoil of 2008/2009, in the UK equities market fundamental volatility has decreased, and liquidity and trading volume have slowly returned. If high-frequency trading (HFT) contributes to volatility, Linton and Atak argue, it might be expected that the ratio of intraday volatility to overnight volatility would have increased as HFT became more commonplace, but Linton and Atak do not find evidence to support that hypothesis. They note that the frequency of large intraday price moves was high during the crisis period, but since the end of 2009 the frequency has declined to more normal levels.

CBT and HFT are relatively new phenomena so the empirical literature examining their role is still nascent. Research thus far provides no direct evidence that HFT has increased volatility.[4] Significant challenges in evaluating HFT are that much of its growth coincides with the 2008/2009 turmoil and the lack of data fully characterising HFT.[5] Indirect studies of CBT

---

[3] Stability can differ from volatility by placing significantly greater weight on large, infrequent price changes.

[4] See DR 12, Brogaard, J. (2011), High Frequency Trading and Its Impact on Market Quality, http://ssrn.com/abstract= 1641387; Chaboud, A., Chiquoine, B., Hjalmarsson, E., & Vega, C. (2009), Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market, http://ssrn.com/abstract=1501135; & Hasbrouck, J., & Saar, G. (2011), Low-Latency Trading, Working paper, Cornell University.

and HFT provide interesting evidence highlighting the importance of further study with better data, but are subject to various interpretations.[6]

It is something of a cliché to say that CBT can lead to "Black Swan" events, i.e. events that are extremely rare but of very high consequence when they do occur. Of course, a more computerised world is more vulnerable to some types of catastrophic events, such as power failures, major solar emissions, cyber-attacks and "outages" of server computers; any one or more of these could in principle lead to system-wide failure events.

However, as far as financial stability is concerned, the more interesting and significant aspects have to do with the general *nonlinear dynamics* of the financial system. Put simply, the dynamics of a system (how it changes over time) are nonlinear if a given change in one variable may either lead to a small change in another variable or to a large change in that other variable, depending on the current level of the first variable. Nonlinear systems can sometimes exhibit very large changes in behaviour as a result of very small alterations of key parameters or variables, and in some cases they are "complex" systems, for which it is impossible to predict the long-term behaviour because an observer can never know the relevant key values with sufficient accuracy. Furthermore, because economic systems are composed of "agents" (individuals, firms, regulators, etc) that interact in various ways, the particular structure of the network of interactions can have a very significant effect on the dynamics of the system too.[7] Currently the scientific literature on complex nonlinear dynamics of networked systems is (in comparison to other fields) in its infancy with regards to concrete predictions and reliably generalisable statements, but can be put on a solid modelling footing by integrating with the foundations of the financial economics literature: this can offer glimpses of important insights, and we rely on those in this report.

Market crashes have been around forever: Zweig, 2011[8] relates the story of a "Flash-Crash" type event in 1962, which the US Securities and Exchange

---

[5] Jovanovic, B., & Menkveld, A. (2011) Middlemen in Limit-Order Markets, http://ssrn.com/abstract=1501135, have a comparison of the volatility of Dutch and Belgian stocks before and after the entry of one HFT firm in the Dutch stocks and find that the relatively volatility of Dutch stocks declines slightly.

[6] For example, Zhang, F. (2010), High-Frequency Trading, Stock Volatility, and Price Discovery, http://ssrn.com/abstract=1691679 , proxies for HFT with a measure of daily trading volume not associated with changes in quarterly institutional holdings. Zhang finds an association between this trading volume measure and "excess" volatility. While the proxy likely relates to CBT, the correlation is difficult to interpret as arising from HFT because the volume-volatility relation appears well before the adoption of HFT as currently defined. Finally, a stronger relation between volume and volatility can result from increased welfare enhancing investor risk sharing. Therefore, indirect studies of HFT and CBT such as this are not recommended as a basis for policy options.

[7] See DR6, DR7 & Haldane, A. & May, R. (2011), Systemic risk in banking ecosystems, *Nature* **469**:351-355.

[8] Zweig, J. (2010), Back to the future: Lessons from the forgotten 'flash crash' of 1962. *The Wall Street Journal*, May 29, 2010. http://online.wsj.com/ article/SB10001424052748703957604575272791511469272.html;

Commission analysed in a report published in 1963[9]. Nevertheless, the problem of understanding mechanisms that underlie system-level events in CBT environments is more recent. A good illustration for the sort of systemic events that mechanical rule-following is able to generate can be found in the portfolio-insurance-led market decline of 1987[10]. In order to hedge their risks, as stock indices dropped, portfolio insurers were required to adjust their 'delta-hedge' a holding of stocks used to balance risk. However, the values of the stocks in the delta-hedge holdings were used to calculate the value of the index. So, as the delta-hedge was adjusted because the index had dropped, stocks were sold, the selling depressed prices, and that pushed the index even lower; this then caused another adjustment of the delta-hedge holdings, which pushed the index even lower still. This positive feedback loop (the effects of a small change looping back on themselves and triggering a bigger change, which again loops back, and so on) had a profoundly damaging effect, leading to major share sell-offs. This loop is illustrated in Figure 1. Such destructive feedback loops can generate nonlinear dynamics and can operate until delta-hedges no longer need to be adjusted or until a market halt is called. The mechanical and unthinking execution of such "program trades" led in 1987 to huge selling pressure and to price falls much deeper than were warranted by actual market conditions.

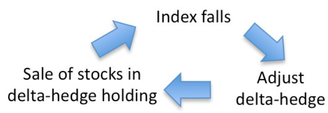


Figure 1: Hedge Feedback Loop

The discrepancy between market prices and true values of securities (given by the fundamental value of the company or resource) when the former are depressed by forced sales in a nonlinear, self-fulfilling and sometimes self-exciting frenzy constitutes a major market instability which comes at a substantial social cost. It might be argued that the more trading decisions are taken by "robot" CBT systems, the higher the risk of such wild feedback loops occurring within a financial system. This "endogenous risk" is the logical thread that runs through much of the rest of this paper. The endogenous risk from human programmed algorithms may differ in important ways from feedback loops and risks in markets with greater direct human involvement.

---

[9] US Securities and Exchange Commission (1963), Report of Special Study of Securities Markets, Chapter 13, available from: http://c0403731.cdn.cloudfiles.rackspacecloud.com/collection/papers/1960/1963_SSMkt_Chapter_13_1.pdf.

[10] See DR9, and also Gennotte, G. & Leland, H. (1990). Market liquidity, hedging, and crashes. *American Economic Review*, **80**:999–1021.

3. How is CBT thought to affect financial stability?

3.1 Mechanisms of instability

It seems unlikely that the future of CBT in the financial markets leads merely to a faster system, and therefore to more frequent crashes and crises, purely on the (metaphorical) basis that the same old movie is now being played at fast-forward speed. Rather, it seems more likely that despite all its benefits, CBT may lead to a qualitatively different and more obviously nonlinear financial system in which crises and critical events are more likely to occur in the first place, even in the absence of larger or more frequent external fundamental shocks. Here some of the insights into what the precise mechanisms could be are outlined.

Three main mechanisms that can lead to instability and losses can be summarised as follows:

- **Sensitivity:** systemic instability can occur if a CBT environment becomes more sensitive to small changes or perturbations. If financial dynamics in a CBT world become sufficiently nonlinear so that widely different outcomes can result from only small changes to one or more current variables (the so-called "butterfly effect" in chaotic dynamics), then the observed prices and quantities are prone to cascades, contagions, instability and path-dependency.[11] The mechanisms of deviation may be the information sensitivity mechanisms, and/or the internal feedback loops, both of which are discussed further below. Even if the effects were temporary and the original driving variables were to revert to their long-term average values over time, some further irreversible events may have occurred, such as financial losses or even bankruptcies due to forced sales or to the triggering of penalty clauses in contracts.

- **Information:** the existence of excessive nonlinear sensitivities can be due to informational issues. Informally, this is concerned with who knows what, when. The information structure of a market has the potential to exacerbate or reduce market swings through a number of subtle, and sometimes contradictory, effects. To illustrate this, academic studies have explored behaviour that arises in coordination games with diffuse information[12]. In these game scenarios, information is diffuse and agents coordinate to create a 'bank-run' on an institution, a security or a currency if a given publicly observed signal is bad enough. Only very small differences in the signal, say the number of write-offs of a bank, determine

---

[11] See DR7 and also Haldane and May (2011) *op. cit.*

[12] See, for example: Carlsson, H. & Van Damme, E. (1993), Global Games and Equilibrium Selection. *Econometrica*, **61**:989-1018. And also Morris, S. & Shin, H. (2002), Global Games: Theory and Applications. In Dewatripont, L. & Turnovsky, S., (eds), *Advances in Economics and Econometrics, the Eighth World Congress*. Cambridge Uni. Press.

whether creditors run or stay. Violent cascades of failures over an entire market system can be triggered by such small events.

- **Endogenous risk:** this term, commonplace in the financial literature[13] identifies features *wholly within* the financial markets that lead in some situations to the sudden (due to the nonlinearities involved) emergence of positive (i.e. mutually reinforcing) and pernicious feedback loops, whether market participants are fully rational or otherwise[14] [15].

In the discussion that follows, a number of feedback loops that can contribute to endogenous risk are explored.

Risk feedback loop. Financial crises typically involve endogenous risk of the following sort. Assume that some financial institutions are hit by a loss that forces them to lower the risk they hold on their books. In order to reduce risk, they need to sell risky securities. Since many institutions hold similar securities, the sale of those assets depresses prices further. When institutions are required to practice "mark-to-market accounting" (where the value of some holding of securities is based on the current market price of those securities), the new lower valuations lead to a further hit to bank capital for all institutions holding the relevant securities, and also to a further increment in general perceived risk. Those two factors in turn force financial institutions to shed yet more of their risks, which in turn depresses prices even further, and so forth. A small initial fundamental shock can lead to disproportionate forced sales and value-destruction because of the amplifying feedback loops hard-wired into the fabric of financial markets.[16] Versions of this loop apply to HFT market makers: given the tight position and risk limits HFT operate under, losses and an increase in risk lead them to reduce their inventories, thereby depressing prices, creating further losses and risk, closing the loop. The value-destruction in turn can lead banks to stop performing their intermediation role with adverse spillover effects on the real economy.

---

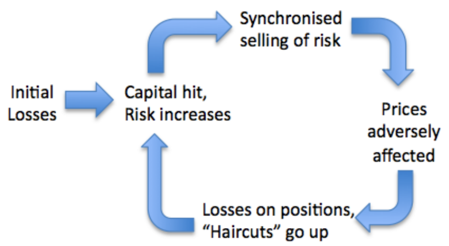
[13] See: Danielsson, J. and Shin, H. (2003). *Endogenous Risk in Modern Risk Management: A History*. Risk Books; Danielsson, J., Shin, H. S., and Zigrand, J.-P. (2010). *Balance sheet capacity and endogenous risk*. Mimeo, www.riskresearch.org; and Shin, H. (2010). *Risk and Liquidity*: 2008 Clarendon Lectures in Finance. Oxford.

[14] See DR2, DR6, DR7, and DR9.

[15] See also M. O'Hara (1995), *Market Microstructure Theory*. Blackwell Publishers.

[16] For further details see Brunnermeier, M. & Pedersen, L. (2009), Market Liquidity and Funding Liquidity, *Review of Financial Studies*, **22**:2201-2238. For empirical evidence on the effect of losses or redemptions see Joshua, C. & Stafford, E. (2007), Asset Fire Sales (and Purchases) in Equity Markets, *Journal of Financial Economics*, **86**:479-512. For evidence on how shocks can propagate across different financial institutions see Khandani, A. & Lo, A. (2007), What Happened To the Quants in August 2007? *Journal of Investment Management* **5**:29-78 and Khandani, A. & Lo, A. (2011), What Happened To the Quants In August 2007? Evidence from Factors and Transactions Data, *Journal of Financial Markets* **14**:1-46.

Volume feedback loop. Whether the official CFTC/SEC report[17] version of the “Flash Crash” events of 6 May 2010 turns out to be accurate and complete or not (see DR4 and DR7 for further discussion), it does illustrate a potential driver of risk. The report outlines a possible scenario whereby some high frequency trader (HFT) algorithms may *directly* create feedback effects via their tendency to hold small positions for short time periods: a “hot-potato” or “pass-the-parcel” dynamic occurred on May 6 where trading amongst HFTs generated very large volumes but the overall net position hardly changed at all: financial instruments were circulating rapidly within the system, and this increase in volume triggered other algorithms which were instructed to sell more aggressively in higher volume markets, presumably on the basis that higher volume means lower market impact, to sell into the falling market, closing the loop[18].



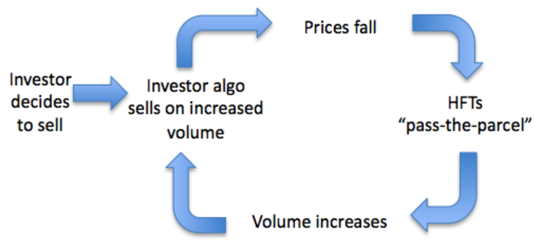


Figure 2: Risk Feedback Loop Figure 3: Volume Feedback Loop

Shallowness feedback loop. Closely related is the potential feedback loop described by Angel (1994)[19] and Zovko and Farmer (2002)[20]. Assume an initial increase in volatility, perhaps due to news. The distribution of bids and asks in the order book adjusts and becomes more dispersed. [21] With everything else constant, incoming *market orders* (i.e., orders to buy or sell at the market’s current best available price) are more able to move the market reference price (based on the most-recent transaction price), and this

---

[17] CFTC & SEC (2010). *Findings regarding the market events of May 6, 2010.* Official report published September 30, 2010 by the Commodities & Futures Trading Commission and the Securities & Exchange Commission.

[18] Kirilenko, A., Kyle, A., Samadi, M. & Tuzun, T. (2011). The Flash Crash: The Impact of High Frequency Trading on an Electronic Market, http://ssrn.com/abstract=1686004 provide evidence that some market participants very rapidly bought and sold from each other with small changes in position over very short horizons during the Flash Crash. While this is a case study of only one (index) security over a few trading days, it highlights the importance of better understanding certain CBT and HFT strategies and their interactions with each other and other market participants. While not able to directly indentify CBT or HFT, Easley *et al.* provide evidence the speed and magnitude of unusual trading behaviour in more securities during the Flash Crash: see Easley, D., Lopez de Prado, M. & O'Hara, M. (2011). The Microstructure of the ‘Flash Crash’: Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading. *The Journal of Portfolio Management*, **37**: 118-128

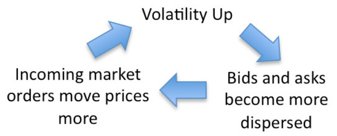
[19] Angel, J. (1994). *Limit versus market orders*. Working Paper, Georgetown University.

[20] Zovko, I. & Farmer, D. (2002). The power of patience: a behavioral regularity in limit-order placement. *Quantitative Finance*, **2**(5):387-392.

[21] For an examination of these types of book changes prior to news about earnings see Lee, C. , Mucklow, B., & Ready, M., (1993). Spreads, depths, and the impact of earnings information: an intraday analysis. *Review of Financial Studies*, **6**:345-374.

increase in volatility in turn feeds back into yet more dispersed quotes, and the loop is closed.

News feedback loop. Many automated HFT systems work primarily on numeric information from market data sources concerning prices and volumes of market orders, but some HFT systems include a *news listener* component that scans headlines for tags and acts upon them immediately by broadcasting the tag to all other components of the HFT system (*news analytics*, the computerised analysis of text-based news and online discussions to generate CBT trading systems, is discussed in greater detail in DR8). HFTs buy or sell depending on where prices are relative to the HFT's own perceived fair value; if the transactions of HFT systems are reported in news feeds, and picked up on by other HFT systems, that can lead those systems to revise their price in a direction that encourages them (or other HFTs) to make similar trades.



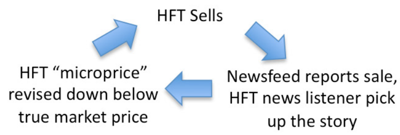


Figure 4: Shallowness Feedback Loop Figure 5: News Feedback Loop

Delay feedback loop. Eric Hunsader of Nanex Corp. suggests[22] the potential of the following very simplified feedback loop that may have operated during the Flash Crash. Consider a fragmented market suffering from overall selling pressure of a basket of high-capitalisation stocks (e.g. originating from the sales of E-Mini futures), and assume that the NYSE quotes lag by a bit. Since the market is falling, the delayed NYSE bids appear to be the most attractive to sellers, and all sales are routed to NYSE, regardless of the fact that actual bids were lower. Algorithmic momentum HFTs short those stocks and given the oddness, HFTs may sell inventories. A second feedback loop then reinforces the first one: as delays creep in and grow, the increased flurry of activity arising from the previous feedback loop can cause further misalignments in bid/ask time stamps, closing the *delay feedback loop* which is amplifying the pricing feedback loop.[23]

Index feedback loop. The CFTC/SEC final report on the Flash Crash argued that the extreme volatility of the individual component securities spilled over into the ETF (exchange-traded fund) markets and led market

---

[22] See DR7 and also: Hunsader, E. (2010). *Analysis of the flash crash, date of event: 20100506, complete text.* Nanex Corp., http://www.nanex.net/20100506/FlashCrashAnalysis CompleteText.html.

[23] Consistent with delay feedback loops having stronger effects in securities where trading is spread across more markets, Madhavan (2011) finds that across stocks, the impact of the Flash Crash is positively related to measures of fragmentation in the month before: Madhavan, A. (2011), *Exchange-Traded Funds, Market Structure and the Flash Crash*, Working Paper, BlackRock, Inc.

makers to pause their market making activities. In return, the illiquid and stub ETF prices for aggregates provide false systematic factor signals, feeding back into the pricing of individual securities, and thereby closing the loop.[24]

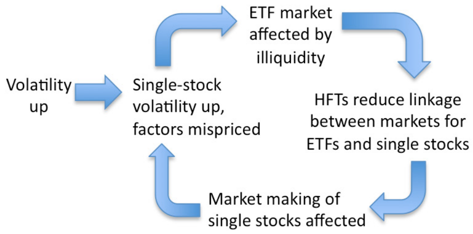


Figure 6: Quote-Delay Feedback Loop Figure 7: Index Feedback Loop

3.2 Interactions between mechanisms

The strength of the feedback loops is tied in with and driven by various variables, especially the capitalisation levels and the leverage ratios of financial institutions as well as the degree of diversity of market participants. For instance, if liquidity is provided by lightly capitalised HFT operators as opposed to deep-pocketed market-makers with large inventories,[25] then the passing-of-the-parcel as well as the resulting feedback loops are stronger because inventory management with little capital requires the quick offloading of positions. In that sense at least, substituting speed for capital works well for market-making purposes in normal times but may not work well in more stressed times where the lack of capital can extremely quickly contribute towards positive instead of negative feedback loops.[26] Similarly, the less diversity that market participants show, the more in unison they act and so the stronger the feedback loops. Moreover, diversity itself can change for the worse during an episode of endogenous risk as an unintended consequence of the combined interactions of risk-management systems, regulatory constraints, margin calls and mark-to-market accounting requirements, which can lead to instantaneous synchronisation of actions among a group of institutions if they are all subject to the same regulations, constraints, and coordination devices. For instance, the CFTC and SEC found that during the crucial minutes of the Flash Crash, liquidity providers switched to becoming liquidity demanders and sold aggressively and in an unsophisticated fashion into a falling market once their inventories reached a certain level. In a situation where liquidity provision is in the hands of a small

---

[24] Madhavan (2011) o*p. cit.* analyses the feedback loop between ETFs and the underlying securities during the Flash Crash.

[25] For empirical evidence on the importance of liquidity providers balance sheets see Comerton-Forde, C., Hendershott, T., Jones, C., Moulton, P., & Seasholes, M. (2010) Time Variation in Liquidity: The Role of Market Maker Inventories and Revenues. *Journal of Finance*, **65**:295-331.

[26] For further details on how intermediaries manage risk and how this affects price dynamics see Duffie, D. (2010), Presidential Address: Asset Price Dynamics with Slow-Moving Capital, *Journal of Finance*, **65**:1237-1267. For empirical measures of the magnitude see Hendershott, T. & Menkveld, A. (2011), *Price Pressures*. http://ssrn.com/abstract=1411943.

number of large but slightly capitalised players, this can lead to a dry-up of liquidity when a small number of these large players switch liquidity provision off (see also DR7 for an estimate of the strength of this effect). In fact informational asymmetries (where one player knows that they know less than another) have the power to strengthen the other two mechanisms in a variety of ways (for more details, see DR9). This is because investors may mistake a temporary forced sale for a sale based on negative inside information, leading to lower prices still.

Similarly, in CBT environments it appears to be much easier for predatory investors to accumulate positions anonymously than it was in times past when all traders interacted physically on the floor of an exchange, partly due to the fact that only anonymous quotes in a limit order book are observed as opposed to a physical person buying contracts in a central pit, and partly due to the ease of using “unlit” or “dark pool” venues where trades are only reported after they have happened. If such players were able to push prices purposefully into cascades, they might be able to generate profitable trades, giving them an incentive to create instability. While such situations can be conceived, access to the data required to establish how easy it is to launch a cascade or indeed how common is this style of predatory behaviour is difficult, assuming it exists at all.

Furthermore, information in CBT environments can present many subtle dangers. For instance, even if all market participants knew a certain event has not occurred, market prices and quantities would still not be able to completely discount the event, for while everyone knows the event has not occurred, it may not be “commonly known” that the event has not occurred in the sense that there can still be participants who do not know that others know that the event has not occurred. It can be argued that technology has removed, to a certain extent, common knowledge from markets. Markets have become networked distributed computing environments (and a well-known theorem states that events cannot be common knowledge in distributed computer environments due to the absence of concurrent centrality of observation, a phenomenon sometimes referred to as the “electronic mail game”)[27]. This has in effect introduced large amounts of complexity, if only due to the fact that the outcome of the market now depends in a non-trivial way on what each trader believes any other trader believes about yet any other traders' beliefs about all other traders, and so forth. The direct link between market outcomes and the fundamental events that ought to act as anchors for valuation has been severed and replaced by a complex web of iterated and nested beliefs.

[27] See: Rubinstein, A. (1989). The electronic mail game: Strategic behavior under almost common knowledge. *American Economic Review*, **79**:385–391; and Halpern, J. & Moses, Y. (1990). Knowledge and common knowledge in a distributed environment. *Communications of the ACM*, **37**(3):549–587.

3.3 Socio-technical factors: normalisation of deviance

Cliff and Northrop propose in DR4 that the "Flash Crash" event in the US financial markets on 6 May 2010 is in fact an instance of what is known technically as "*normal failure*", and they explain that such failures have previously been identified in other complex engineered systems. They argue that major systemic failures in the financial markets, at a national or global scale, can be expected in future, unless appropriate steps are taken.

Normal failures (a phrase coined in 1984 by Charles Perrow[28]) in engineered systems are major system-level failures that become almost certain as the complexity and interconnectedness of the system increases. Previous examples of normal failure include the accident that crippled the *Apollo 13* moon mission, the nuclear-power accidents at Three Mile Island and Chernobyl, and the losses of the two US space-shuttles, *Challenger* and *Columbia*.

As Cliff and Northrop note in DR4, the American sociologist Diane Vaughan has produced detailed analyses of the process that gave rise to the normal-failure losses of *Challenger* and *Columbia*, and has argued that the key factor is the natural human tendency to engage in a process that she named[29] *"normalisation of deviance"*: when some deviant event occurs that was previously thought to be highly likely to lead to a disastrous failure, if it then happens that actually no disaster does occur, there is a tendency to revise the agreed opinion on the danger posed by the deviant event, assuming that in fact it is normal: the "deviance" becomes "normalised". In essence, the fact that no disaster has yet occurred is taken as evidence that no disaster is likely if the same circumstances occur again in future. This line of reasoning is only broken when a disaster does occur, confirming the original assessment of the threat posed by the deviant event.

Cliff and Northrop argue that the Flash Crash was, at least in part, a result of normalisation of deviance. For many years, long before 6 May 2010, concerns about systemic effects of rapid increases in the price volatility of various instruments had led several UK exchanges to implement "circuit breaker" rules, requiring that trading in a security be suspended for some period of time if the price of that security moved by more than some percentage within a sufficiently short time-period. In response to the Flash Crash, the USA's SEC has now enforced similar mechanisms in the US markets with the aim of preventing such an event re-occurring. Thus, it seems plausible to argue that before the Flash Crash occurred there had been a significant degree of normalisation of deviance: high-speed changes in the prices of equities had been observed, market participants were well

---

[28] Perrow, C. (1984). *Normal Accidents: Living with High-Risk Technologies*. New York: Basic Books.

[29] Vaughan, D. (1997), *The Challenger Launch Decision: Risky Technology, Culture and Deviance at NASA.* University of Chicago Press.

aware that that could lead to a high speed crash, but these warning signals were ignored and the introduction of safety measures that could have prevented them was resisted.

Moreover, it could plausibly be argued that normalisation of deviance has continued to take place in the markets since the Flash Crash. There are anecdotal stories (summarised in DR4) that the speed of price fluctuations occurring within the limits of circuit breaker thresholds seems to be increasing in some markets; and there is evidence to suggest that another flash crash was "dodged" on 1 September 2010, in a similarly bizarre period when quote volumes exceeded even those seen at peak activity on 6 May 2010, but no official investigation was commissioned to understand that latter event. Furthermore, the circuit-breaker mechanisms in each of the world's major trading hubs are not harmonised, exposing arbitrage opportunities for exploiting differences; computer and telecommunications systems can still fail, or be sabotaged by enemies of the system, and the systemic effects of those failures may not have been fully thought through.

Of course, the next Flash Crash will not be exactly the same as the last one, the SEC's new circuit breakers will probably see to that. But there are no guarantees that another event, just as unprecedented, just as severe, and just as fast (or faster) than the Flash Crash cannot happen in future. Normalisation of deviance can be a very deep-running, pernicious process. After Challenger, NASA addressed the immediate cause (failure of a seal on the booster rockets), and believed the Shuttle to be safe. That was no help to the crew of Columbia.

Reassurances from regulators are likely to sound somewhat hollow for as long as people can remember the near-total failure of the regulatory bodies to have anything useful to say about the subprime crisis until shortly after its severity was clear to even the most casual of observers. Light-touch regulation and its consequence for financial markets in the UK were discussed in the 2009 Turner Review[30]. Relatedly, the contributions of the Basel II regulations to the causes of the subprime crisis have been discussed in Shin (2010). The next market failure may well be a failure of risky technology that, like the Flash Crash, has no clear precedent.

Cliff and Northrop argue in DR4 that normalisation of deviance poses a threat to stability in the technology-enabled global financial markets, and that the dangers posed by normalisation of deviance and normal failures are if anything heightened in the financial markets because the globally interconnected network of human and computer traders is what is known in the academic literature as a *socio-technical system-of-systems* (i.e., an interconnected mesh of people and adaptive computer systems interacting with one another, where the global system is composed of constituent entities that are themselves entire independent systems, with no single overall management or coordination). Such systems are so radically different

---

[30] http://www.fsa.gov.uk/pubs/other/turner_review.pdf

from traditional engineered systems that there is very little established science or engineering teaching that allows us to understand how to manage and control such super-systems. This issue of normalisation of deviance in the financial markets and its role in informing the discussion of possible regulatory options was recently discussed in more detail by Haldane (2011)[31].

4. How can CBT affect financial stability in future?

Left to its own devices, the extent of CBT may still grow, and the endogenous risk factors outlined above continue to apply. The mix between the three mechanisms however may have started to find its level. For instance, there are natural bounds on the extent of trading that can be generated by proprietary ("short-termist") HFTs. First, those trades that have such HFTs on both sides of a transaction can generate profits only for one. Secondly, the *no-trade theorem*[32] would predict that once it becomes known that the only trades put on are those by short-termist proprietary traders who do not have an incentive to hold the securities for fundamental reasons, trade will collapse. And lastly, much of the profits and rents of HFTs are competed away under increasing competition. Recent reports suggest that profits of HFT companies have turned downwards[33], and a recent academic study[34] has established that the total profits available for extraction by HFT may not be as large as some people suspect. Looking at trading patterns, there is preliminary evidence that HFT may have reached its equilibrium penetration into London and EuroNext equity trading (see DR5). In a nutshell, CBT may gain market share as more buy-side investors use it, but proprietary intermediation trading is naturally limited by the fundamental trading volume of real-money investors.

If CBT further imposes itself while open-outcry trading pits (where human traders interact in close proximity to each other) and other centralised mechanisms vanish, then the disappearance of common knowledge and the creation of complex belief networks will progress further. On the Chicago Mercantile Exchange for instance, some contracts moved from open-outcry to a hybrid model with open-outcry during waking hours and electronic markets outside of those hours. Over the next five to 10 years, one may see further substitution of floor based trading in favour of CBT, rendering all the effects described above more salient still.

---

[31] Haldane, A.S. (2011). *The Race to Zero*. http://www.bankofengland.co.uk/publications/speeches/2011/speech509.pdf.

[32] Milgrom, P. and Stokey, N. (1982). Information, trade and common knowledge. *J. Economic Theory,* **26**:17–27.

[33] See, e.g., Cave, T. (2011): http://www.efinancialnews.com/story/2010-10-18/high-frequency-party-over.

[34] Kearns, M., Kulesza, A., & Nevmyvaka, Y (2011). Empirical Limitations of High Frequency Trading Profitability. *The Journal of Trading,* **5**(4):50-62. http://ssrn.com/abstract=1678758.

Financial institutions optimise, subject to the regulatory environment, which means that constraints will often be binding and may influence market dynamics in unexpected and often detrimental ways. For instance, the academic literature has identified many circumstances where the "fallacy of composition" appears: the market system is unstable despite the fact that each algorithm in isolation is stable.[35] This strongly suggests that if a testing facility for algorithms was introduced, individual safety is not a sufficient or indeed even a necessary criterion for systemic stability. It follows that in order to predict the future of CBT and financial stability one needs to make assumptions as to the future regulatory and other constraints imposed upon markets and think the new dynamics through carefully. For instance, further in-depth studies may provide indications as to how minimum resting times or minimum tick sizes affect nonlinear market dynamics and financial stability.

A second institutional feature that will matter to future stability is the market segmentation between the various competing trading venues. Aligning valuations for single securities, baskets of securities, and derivatives across venues is a socially useful task that HFTs currently perform for a fee. Social welfare requires that this role be filled in a permanent and well-capitalised fashion. There is a small risk that HFTs, under situations of market stress, margin calls or collateral pressure, will not be able or willing to perform this job.

Non-linearities in liquidity provision (leading to quick reversals between feast and famine) are an important root cause of system-wide non-linear dynamics that deserve further study. Most of the time HFT adds to liquidity, but some of the time (in periods of stress or crisis) it subtracts liquidity, causing price discontinuities. These liquidity non-linearities have probably become more acute in a HFT world because of the effects discussed above, which have made the "market makers problem" of inventory and information management not only different but also altogether more difficult.

Finally, in closing, in a world with multiple trading and pricing venues that are interconnected by HFTs, the network topology determines the stability and the flow of information and trades. With the predicted proliferation of company-owned "dark pools", the aggregate liquidity across all venues may well be larger than with single monopolised exchanges, but the dynamic behaviour of liquidity will depend more and more on the network structure as well as on the specifics of the HFTs which link the trading venues. A liquidity shock on one venue that might have gone unnoticed if there was one large centralised exchange can now affect prices on that venue. In normal times, the aberrant price would quickly disappear as cross-trading-venue HFTs buy low and sell high. But in stressed times, the capital of HFTs may be limited, or the HFTs themselves may start to doubt the prices (as happened during the Flash Crash) and refrain from arbitraging. Real-money investors then start to mistrust valuations across the board, and the resulting pressures mean that HFTs no longer contribute to liquidity provision, which makes

[35] This fallacy is discussed in Samuelson, P. (1947). *Foundations of Economic Analysis*, Harvard University Press.

price divergence across trading venues worse still. And so the shock is transmitted through the network, and its effects are reinforced by positive feedback, as illustrated in Figure 8. Trades and transactions will happen at socially inefficient prices, and mark-to-market valuations can only be done to multiple and illiquid marks. Understanding how to avoid such situations, and to contain them when they do occur, is a topic for further research.

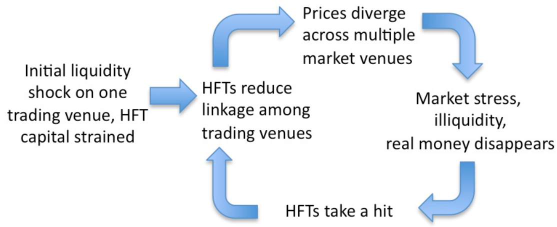


Figure 8: Systemic Divergence Feedback Loop

5. Conclusion

Markets with significant proportions of computer based high frequency traders are a recent phenomenon. One of the most novel aspects of their dynamics is that interactions take place at a pace where human intervention could not prevent them, and in that sense an important speed limit has been breached. Because of the speed advantages that computers can offer in comparison to humans, computer based trading is now almost obligatory. This gives rise to the potential for new system-wide phenomena and uncertainties. One key issue is that information asymmetries become more acute (and indeed different in nature) than in the past; and the primary source of liquidity provision has changed, to computer based and high-frequency trading systems, which has implications for the robustness of markets in times of stress.

Research thus far provides no direct evidence that high frequency computer based trading has increased volatility. But, in certain specific circumstances, self-reinforcing feedback loops within well-intentioned management and control processes can amplify internal risks and lead to undesired interactions and outcomes. These feedback loops can involve risk-management systems, and can be driven by changes in market volume or volatility, by market news, and by delays in distributing reference data. A second cause of market instability is social: *normalisation of deviance*, a process recognised as a major threat in the engineering of safety-critical systems such aeroplanes and spacecraft, can also affect the engineering of computer based trading systems.

Paper 2: The impact of computer trading on liquidity, price efficiency/discovery and transaction costs

Oliver Linton, Cambridge University

Maureen O'Hara, Cornell University

Key findings

Overall, liquidity has improved, transaction costs are lower, and market efficiency has not been harmed by computerised trading in regular market conditions.

The nature of market making has changed, shifting from designated providers to opportunistic traders. High frequency traders now provide the bulk of liquidity, but their use of limited capital combined with ultra-fast speed creates the potential for periodic illiquidity.

Computer –driven portfolio rebalancing and deterministic algorithms create predictability in order flows. This allows greater market efficiency, but also new forms of market manipulation.

Technological advances in extracting news will generate more demand for high frequency trading, while increased participation in this will limit its profitability.

Executive summary

Computerised trading has become the norm in markets, affecting everything from portfolio selection to order submission to market making to clearing and settlement. New players such as high frequency traders (HFT) and new strategies employing algorithmic trading (AT) are influencing the behaviour and quality of markets. With estimates of HFT participation rates in European equity markets of 30-50% and in US equity markets as high as 70%, HFT has a profound influence on current markets.

A natural concern is how such computerised trading is affecting the quality of markets. Market quality is usually expressed in terms of liquidity, transaction costs, and price efficiency. This study provides evidence on these effects in current markets and discusses the likely future impacts of computerised trading on market quality.

The evidence suggests that computerised trading (whether in the guise of high frequency trading or algorithmic trading) has generally improved market quality. Liquidity, as measured by bid/ask spreads and other metrics, has improved over the last decade. During this period, transaction costs have also fallen for both retail and institutional traders. These liquidity and transaction cost effects have been particularly pronounced for large stocks. There is also evidence that market prices are more efficient, consistent with the hypothesis that computerised trading links markets and thereby facilitates price discovery.

While overall liquidity has improved, there appears to be greater potential for periodic illiquidity. The nature of market making has changed, with high frequency traders now providing the bulk of such activity in both futures and equities. Unlike traditional designated specialists, high frequency traders typically operate with little capital, hold small inventory positions, and have no obligations to provide liquidity during periods of market stress. The speed of trading as well as the interconnectedness of markets made possible by HFT can transmit disruptions almost instantaneously across markets. The US Flash Crash, as well as more recent smaller illiquidity events, illustrates this increased potential for periodic illiquidity.

The predictability of order flows arising from mutual fund rebalancing and deterministic trading algorithms creates potential disruptions to market quality. New forms of manipulation, such as algorithms programmed to take advantage of other algorithms, can raise trading costs and move prices away from efficient levels. Increasing competition in the high frequency space should limit such effects, as would new regulations.

While there is some evidence that HFT decreased in early 2011, it has re-emerged with the recent volatility in the market. HFT and algorithmic trading will remain important going forward as increased technological advances provide profitable opportunities to trade on computerised news analysis (NA) techniques. New regulatory changes may limit the profitability of HFT in the future.

1. Introduction

Technology has transformed asset markets, affecting the trading process from the point of asset selection all the way through to the clearing and processing of trades. Portfolio managers now use computerised order management systems to track positions and determine their desired trades, and then turn to computerised execution management systems to send their orders to venues far and wide. Computer algorithms (AT), programmed to meet particular trading desires, slice and dice orders to trade both temporally across the trading day and spatially across markets. High frequency traders (HFT) use ultra-fast computers and market linkages both to make and take liquidity across and between markets. Transaction cost analysis, using computers to capture price movements in and across markets, then allows asset managers to calculate their trade-specific transaction costs for particular trading strategies, and predict their costs from using alternative strategies.

What is particularly striking is that virtually all of these innovations have occurred within the past 10 years.[36] In this short interval the market ecology has changed, with markets evolving from traditional exchange-based monopolies to networks of computer-linked venues.[37] Yet, while the process of trading in markets has changed, the function of markets remains the same: markets provide liquidity and price discovery that facilitates the allocation of capital and the management of risk. The purpose of this paper is to highlight what is known, and unknown, about the impact of computer trading on liquidity, transaction costs and market efficiency. The likely future evolution of computer trading on these dimensions of market quality is also considered.

2. The impact of computer trading on liquidity, transaction costs, and price efficiency

Determining the impact of technology on market quality (the general term used to describe the liquidity, transaction cost, and price efficiency of a market) is complicated by the many ways in which computers affect the trading process. Moreover, there is not even complete agreement on how to define some of these technological innovations. High frequency trading is a case in point. HFT was virtually unknown five years ago, yet high frequency traders at times participate in 70% or more of trades in equities and futures markets. The U. S. Securities and Exchange Commission (SEC) 2010

[36] See DR5 for discussion of the origins and growth of computerised trading.

[37] See DR6 for discussion.

Concept Release on Equity Market Structure (SEC (2010)) describes HFT as employing technology and algorithms to capitalise on very short-lived information gleaned from publicly available data using sophisticated statistical, econometric, machine learning, and other quantitative techniques. Yet, even within this general description, the SEC notes the difficulty in characterising what high frequency trading actually means:

"The term is relatively new and is not yet clearly defined. It typically is used to refer to professional traders acting in a proprietary capacity that engage in strategies that generate a large number of trades on a daily basis… Other characteristics often attributed to proprietary firms engaged in HFT are: (1) the use of extraordinarily high-speed and sophisticated computer programs for generating, routing, and executing orders; (2) use of co-location services and individual data feeds offered by exchanges and others to minimize network and other types of latencies; (3) very short time-frames for establishing and liquidating positions; (4) the submission of numerous orders that are cancelled shortly after submission; and (5) ending the trading day in as close to a flat position as possible (that is, not carrying significant, unhedged positions over-night)."[38]

Despite the lack of clarity as to the exact meaning of HFT, there is little disagreement as to its importance in markets. Many high frequency traders act as market makers by placing passive limit orders onto electronic order books.[39] These passive orders provide the counter-party for traders wishing to find a buyer or seller in the market. In addition, high frequency traders often engage in statistical arbitrage, using their knowledge of correlations between and within markets to buy an asset trading at a low price and simultaneously sell a correlated asset trading at a higher price. This activity essentially "moves" liquidity between markets, providing a new dimension to the market making function. The centrality of this role means that HFT are involved in a large percentage of market volume. Estimates of HFT involvement in US equity trading are as high as 77%, with estimates of HFT involvement in European equities markets ranging from 30-50%. Estimates of HFT in futures and FX markets are in a similar range. Tabb Securities estimates profits of high frequency traders in 2010 at $7.2 billion, although Kearns, Kulesza, and Nevmyvaka (2010) argue that the actual number is much lower.[40]

There are some important differences between such high frequency market making and its traditional specialist-based counterpart. HF market makers rely on high speed computer linkages (often achieved by co-locating their

---

[38] See Securities and Exchange Commission, 2010, Concept Release on Equity Market Structure, Release No. 34-61458; File No. S7-02-10, page 45.

[39] See Brogaard (DR10) and Hendershott (DR12).

[40] M. Kearns, A. Kulesza, and Y. Nevmyvaka, 2010, Empirical Limitations of High Frequency Profitability, Working Paper, University of Pennsylvania.

computers at the exchange) to enter massive numbers of trades with the goal of earning the bid-ask spread. Such traders generally hold positions for very short periods of time (in some cases, micro-seconds) and some operate with very low levels of capital. Whereas, in traditional markets, specialists had obligations to stand ready to buy and sell, HF market makers trade opportunistically: they typically do not hold large inventory positions and they manage their risks by curtailing trading when market conditions are too adverse.[41] This behaviour raises the spectre of periodic illiquidity.

The debate surrounding high frequency trading has become increasingly heated, reflecting the varied perspectives on the ability (and desirability) of high frequency traders to move faster (and on the basis of potentially greater information) than other traders.[42] Paul Krugman represents the contrarian view of high frequency trading:

"It's hard to imagine a better illustration (of social uselessness) than high-frequency trading. The stock market is supposed to allocate capital to its most productive uses, for example by helping companies with good ideas raise money. But it's hard to see how traders who place their orders one-thirtieth of a second faster than anyone else do anything to improve that social function... we've become a society in which the big bucks go to bad actors, a society that lavishly rewards those that make us poorer."[43]

Rhetoric aside, while the issues surrounding computer based trading, and HFT in particular, are complex, they are amenable to economic analysis. A useful starting point is to consider how market quality has fared as this new market ecology has developed.[44]

---

[41] It should be noted that traditional specialists also typically avoided holding large inventory positions. The number of specialists obliged to make markets, meanwhile, has dwindled in the face of new competition. On the London Stock Exchange (LSE) there are official market makers for many securities (but not for shares in the largest and most heavily traded companies, which instead use an automated system called TradElect). Some of the LSE's member firms take on the obligation of always making a two-way price in each of the stocks in which they make markets. Their prices are the ones displayed on the Stock Exchange Automated Quotation (SEAQ) system and it is they who generally deal with brokers buying or selling stock on behalf of clients.

[42] High frequency traders generally use proprietary data feeds to get information on the state of the market as quickly as possible. In the U.S., this means that such traders receive information before it is delivered over the consolidated tape, raising issues of fairness and potential harmful effects on the cost of capital (see SEC [2010] (op cit) for discussion).

[43] New York Times, August 2, 2009.

[44] The development of the innovations discussed here occurred largely in the past decade, a time period also characterised by a very large financial and banking crisis and now a sovereign debt crisis. Moreover, both Europe and the U.S. saw dramatic regulatory change in the guise of MiFID and Reg NMS, respectively. Consequently, care must be taken before ascribing all change in the market's behavior to particular technological innovations.

3. Past: what has been the impact of computer trading on market quality in recent years?

3.1 Liquidity

Liquidity is a fundamental property of a well-functioning market, and lack of liquidity is generally at the heart of many financial crises and disasters. Defining liquidity, however, is problematic. At its simplest level, a market is liquid if a trader can buy or sell without greatly affecting the price of the asset. This simple statement, however, begs a variety of questions. Does it matter how much a trader wants to trade? Doesn't time, or how long it takes to execute a trade, also matter? Won't liquidity depend in part on the trading strategy employed? Might not liquidity mean different things to different traders?

Academics and market practitioners have developed a variety of approaches to measure liquidity. Academics argue that liquidity is best measured or defined by attributes such as tightness, resilience, and depth.[45] Tightness is the difference between trade price and original price. Depth corresponds to the volume that can be traded at the current price level. Resilience refers to the speed with which the price returns to its original level after some (random) transactions. In practice, researchers and practitioners rely on a variety of measures to capture liquidity. These include bid-ask spreads (tightness), the number of orders resting on the order book (depth), and the price impact of trades (resilience). Transaction cost analytical models (TCA) use measures such as realised spreads and effective spreads to measure actual liquidity, and rates of trade and order book dynamics to forecast expected liquidity.

In current markets, a common complaint is that liquidity is transient, meaning that orders are placed and cancelled within a very short time frame and so are not available to the average investor. The counterpoint to this argument is that algorithms now split large orders (the so-called "parent" order) into smaller "child" orders that are executed over time and location. Like the unseen parent order, these child orders are often not totally displayed to the market. So liquidity per se is more of a moving target, and traders seek it out using various computer-driven strategies. A variety of algorithms, such as Credit Suisse "Guerrilla", Goldman Sachs "Stealth", or ITGs "Dark", are designed to find liquidity without revealing the trading intentions, or even the existence, of the order submitter. This dichotomy between the lit and dark markets adds another dimension to the challenge of finding and accessing liquidity in computer-driven markets.

The main question of interest is whether computerised trading (either in the guise of algorithmic trading or high frequency activity more generally) is associated with a decrease or increase in liquidity during regular market

---

[45] See Kyle, A. P., 1985, Continuous Auctions and Insider Trading, Econometrica, 53, 1315-1335, and O'Hara, M., 1995, Market Microstructure Theory, (Blackwell, Oxford) for discussion.

conditions. An equally important question relates to whether such trading exacerbates liquidity problems in situations of market stress.

There are a variety of studies that try to identify computerised trading and its consequences on the order book and transactions. Hendershott, Jones, and Menkveld (2011) use the automation of the NYSE quote dissemination as an implicit experiment to measure the causal effect of algorithmic trading on liquidity.[46] In 2003, the NYSE began to phase in the auto-quote system, which empowered computerised trading, initially for six large active stocks and then slowly over the next five months to all stocks on NYSE. They find that this change narrowed bid-ask spreads which they interpreted as more algorithmic trading improving liquidity and reducing adverse selection. Their evidence is strongest for large stocks. Chaboud et al. (2009) also report results on liquidity in the Electronic Broking Services (EBS) exchange rate market.[47] They find that even though some algorithmic traders appear to restrict their activity in the minute following macroeconomic data releases, algorithmic traders increase their supply of liquidity over the hour following each release.

Hasbrouck and Saar (2010) investigate order book data from NASDAQ during the trading months of October 2007 and June 2008.[48] Looking at 500 of the largest firms, they construct a measure of HFT activity by identifying "strategic runs," which are linked submissions, cancellations, and executions that are likely to be parts of a dynamic strategy. Their conclusion is that increased low-latency activity improves traditional market quality measures such as spreads and displayed depth in the limit order book, as well as reduces short-term volatility.

Brogaard (2010) also investigates the effect of high frequency trading on market quality.[49] He finds that HFTs participate in 77% of all trades and that they tend to engage in a price-reversal strategy. He finds no evidence to suggest that HFTs withdraw from markets in bad times or engage in abnormal front-running of large non-HFT trades. HFTs demand liquidity for 50.4% of all trades and supply liquidity for 51.4% of all trades. HFTs also provide the best quotes approximately 50% of the time.

Turning to Europe, Menkveld (2011) studies in some detail the entry of a new high frequency trader into trading on Dutch stocks at Euronext and a

---

[46] Hendershott,T., C. Jones and A. Menkveld, 2011, Does Algorithmic Trading Improve Liquidity, Journal of Finance, 66, 1-33.

[47] Chaboud, A., B. Chiquoine, E. Hjalmarsson, and C. Vega, 2009, Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market, International Finance Discussion Paper No. 980, Board of Governors of the Federal Reserve System.

[48] Hasbrouck, J. and G. Saar, 2010, Low Latency Trading, Working Paper, Cornell University.

[49] Brogaard, J. ,2010, High Frequency Trading and Its Impact on Market Quality, Working paper, Northwestern University.

new market Chi-X in 2007 and 2008. He shows that the inventory of the high frequency trader ends the day close to zero but varies throughout the day, which is consistent with the SEC definition of HFT. He finds that the high frequency trader's earnings all arise from passive orders (liquidity supply). He also finds that the bid-ask spreads were reduced by a factor of about 30% within a year when comparing with Belgian stocks that were not traded by the HFT entrant.

There are also studies reporting trends in liquidity without specifically linking it to AT/HFT. Castura et al. (2010) investigate trends in bid-ask spreads on the Russell 1000 and 2000 stocks over the period 2006 to 2010.[50] They show that bid-ask spreads have declined over this period and that available liquidity (defined as the value available to buy and sell at the inside bid and ask) improved over time. Angel, Harris, and Spatt (2010) show a slow decrease in the average spread for S&P500 stocks over the period 2003-2010 (subject to some short term up-side fluctuations in 2007/2008).[51] They also find that depth has increased slowly over the relevant period. The evidence also shows that both the number of quotes per minute and the cancellation to execution ratio have increased, while market order execution speed has increased considerably. O'Hara and Ye (2011) examine liquidity issues in the context of how fragmentation has affected market quality.[52] They find that stocks with more fragmented trading had lower spreads and faster execution times. As HFT is more prevalent in the newer venues such as BATS and Chi-X, this evidence is suggestive that HFT is correlated with lower spreads and faster execution speeds.

Two commissioned studies provide evidence for UK markets. Friederich and Payne (DR5) compare the operation of HFT in equities and FX. They find that penetration of algorithmic, dynamic agency flow (i.e. best execution of trades on behalf of clients) on multilateral order books in FX is small relative to equities, perhaps because FX is more liquid and therefore orders do not need to be broken up. They report no trend in volume (the traded value) of FTSE100 stocks traded between 2006 and 2011, but find that bid-ask spreads have decreased while depth has increased. The number of trades, on the other hand, has increased more than five times over this period, implying that the average trade size is now only 20% of its former level. For small UK stocks there are different results. First, the average trade size has not changed so much over the period 2006 to 2011, which suggests that HFT is not so actively involved in their trading. Second, there has been little improvement in the liquidity of small cap stocks.

---

[50] Castura, J., R. Litzenberger, R. Gorelick , and Y. Dwivedi , 2010, Market Efficiency and Microstructure Evolution in U.S. Equity Markets: A High-Frequency Perspective, Working paper, RGM Advisors, LLC.

[51] Angel, J., L. Harris, and C. Spatt, 2010, Equity Trading in the 21st Century, SSRN.

[52] O'Hara, M., and M. Ye, 2011, Is Market Fragmentation Harming Market Quality? , Journal of Financial Economics, 100, 459-474.

Linton (DR1) measures the daily liquidity of the FTSE-All Share index using a low frequency measure, the absolute return to unit of volume. He finds this measure of liquidity has varied considerably over the last ten years, first declining and then rising during the financial crisis and then falling again. The same is true of traded volume. The process driving traded volume of UK equities seems to be highly persistent, which means that bad shocks to volume, like that which occurred in 2008/2009, can take a long time to correct. The main conclusion from this is that "Events" are driving the liquidity of UK equity markets, and that if HFT has a role to play in this trend it is relatively small and insignificant in comparison with the big picture of sovereign debt.

In summary, most evidence points to improvement in liquidity in financial markets as a result of computerised trading and HFT, but that there may be some issues in times of market stress.

3.2 Transaction costs

Trading with computers is cheaper than trading with humans, so transaction costs have fallen steadily in recent years as a result of the automation of markets. Jones (2002) reports the average relative one-way costs paid for trading Dow Jones stocks between 1935 and 2000.[53] He finds the total cost of trading has fallen dramatically in the period 1975- 2000. Angel et al. (2010) show that average retail commissions in the USA have decreased between 2003 and 2010, a period more relevant for inferring the effects of computer trading. They also make a cross country comparison of trading costs as of the end of 2009. According to this study, the United States large capitalisation stocks are the cheapest to trade in the world with a roughly 40 basis point cost. Incidentally, the UK fared quite poorly in this comparison, with an average 90 basis point cost that was worse than the rest of Europe and Canada and only marginally better than emerging economy stocks.

Menkveld (DR16) argues that new entry, often designed to accommodate HFT, had profound effects on transaction costs. For example, the entry of Chi-X into the market for Dutch index stocks had an immediate and substantial effect on trading fees for investors, first through the lower fees that Chi-X charged and then through the consequent reduction in fees that Euronext offered. The strongest effect however was a reduction in clearing fees. A new clearing house entered, EMCF, and this triggered a price war that ended up with a 50% reduction in clearing fees.

In summary, the evidence is that transaction costs have declined in the last decade, mostly due to changes in the trading market structure (which is related closely to the development in HFT).

---

[53] Jones, C., 2002, A Century of Stock Market Liquidity and Trading Costs, Working paper, SSRN.

3.3 Price efficiency

A central claim of financial economists is that more efficient prices (better reflecting fundamental values) in financial markets contribute to more informed financing and investment decisions and ultimately to better allocation of resources in the wider economy. The usual way of measuring efficiency is through the predictability of prices based on certain information. In practice, widely-used measures such as variance ratios and autocorrelation coefficients estimate the predictability of prices based on linear rules.

Several studies in this research project address the question of whether the strategies employed by HFT are likely to improve or worsen price efficiency. Brogaard (DR10) describes how high frequency traders (HFT) make money and how their activities may affect price discovery, for example by making the prices of assets with similar payoffs move more closely together. The profits of HFT come from a variety of sources including passive market making activity, liquidity rebates given by exchanges to reward supplying liquidity, and statistical pattern detection used in so-call "stat-arb" strategies. Greater market making activity should improve efficiency. HFT strategies that enforce the law of one price across assets and across trading venues similarly improve the quality of prices facing investors.[54] Farmer (DR6) cautions that as market ecologies change, the transition to greater efficiency may be slow.

Negative effects on efficiency can arise if HF traders pursue market manipulation strategies. Strategies such as front running, quote stuffing (placing and then immediately cancelling orders), and layering (using hidden orders on one side and visible orders on the other) can be used to manipulate prices. For example, deterministic algorithmic trading such as VWAP (volume weighted average price) strategies can be front-run by other algorithms programmed to recognise such trading. Momentum ignition strategies, which essentially induce algorithms to compete with other algorithms, can push prices away from fundamental values. However, it is clear that price efficiency-reducing strategies like manipulative directional strategies are more difficult to implement effectively if there are many firms following the same strategies. Thus, the more competitive the HFT industry, the more efficient will be the markets in which they work.

There is a variety of evidence suggesting that price efficiency has generally improved in the presence of computer trading. Castura et al. (2010) investigate trends in market efficiency in Russell 1000/2000 stocks over the period 1 January 2006 to 31 December 2009.[55] Based on evidence from variance ratios, they argue that markets become more efficient in the

---

[54] Hendershott (DR12) describes the meaning of price efficiency in the context of high speed markets, and presents the arguments why HFT may improve market efficiency by enabling price discovery through information dissemination.

[55] Castura el al (2010), op cit.

presence of and increasing penetration by HFT. Brogaard (2010) provides further evidence on this question. He estimates that the 26 HFT firms in his sample earn approximately $3 billion in profits annually. Were HFTs not part of the market, he estimates a trade of 100 shares would result in a price movement of $.013 more than it currently does, while a trade of 1,000 shares would cause the price to move an additional $.056. He also shows that HFT trades and quotes contribute more to price discovery than do non-HFT activity.

Linton (DR1) provides evidence based on daily UK equity data (FTSEAllshare). Specifically, he computes variance ratio tests and measures of linear predictability for each year from 2000-2010. The measures of predictability (inefficiency) fluctuate around zero with sometimes more and sometimes less statistical significance. During the financial crisis there was more pronounced inefficiency, but that has since declined. He finds no trend in efficiency in the UK market, whether good or bad.

In summary, the preponderance of evidence suggests that HFT has not harmed, and may have improved, price efficiency.

4. Present: what has been the impact of computer trading on market quality at present?

The Flash Crash in US markets has brought increased scrutiny to the role of episodic illiquidity in markets and its relation to the current computer based market structure. The events of 6 May 2010 have now been extensively documented in two reports by the CFTC (Commodity Futures Trading Commission) and SEC (Securities and Exchange Commission) staff. These reports show that a complex interaction of forces led to the Dow Jones Industrial Average falling 998 points, the largest intra-day drop in US market history. While the Flash Crash lasted less than 30 minutes, for a brief interval more than $1 trillion in market capitalisation was lost. In the aftermath of the crash, more than 20,000 trades were cancelled. A more lasting effect has been an almost continual withdrawal from equity investing by retail traders.

The CFTC-SEC reports highlight the role played by a large algorithmic sell trade in the S&P e-mini futures contract that coincided with the beginning of the crash. While clearly an important factor, the reports also highlight the roles played by a variety of factors such as routing rules, quoting conventions, internalises (the name given to banks and broker/dealer firms that clear order flow internally), high frequency traders, and trading halts. These reports make clear two compelling facts about the current market structure: episodic illiquidity can arise, and when it does, it is rapidly transmitted to correlated markets. That the Flash Crash began in what is generally viewed as one of the most liquid futures contracts in the world only underscores the potential fragility of the current market structure.

A variety of research considers how computer trading may be a factor in precipitating periodic illiquidity. Leland (DR9) highlights the role that forced selling can have on market liquidity. Such selling can arise from various trading strategies, and its effects are exacerbated by leverage. Leland argues that algorithmic trading strategies are triggered in response to automatic data feeds and so have the potential to lead to cascades in market prices as selling triggers price moves that trigger additional selling. Leland also argues that due to forced selling the crash of 19 October 1987 and the Flash Crash have great similarities. As modern HFT did not exist in 1987, this result underscores that market illiquidity is not a new event. What matters for our inquiry is whether current computer-driven practices are causing greater illiquidity risk. Madhavan (2011) argues that fragmentation linked to high frequency trading may be one such cause.[56] He shows that fragmentation measured using quote changes, which he argues is reflective of high frequency activity, has high explanatory power with respect to cross-sectional effects on equity instruments in the Flash Crash.

Kirilenko et al. (2011) provide a detailed analysis of high frequency futures traders during the Flash Crash.[57] They found that high frequency traders initially acted as liquidity providers but that as prices crashed some HFT withdrew from the market while others turned into liquidity demanders. They conclude that high frequency traders did not trigger the Flash Crash but their responses to unusually large selling pressure increased market volatility. Easley, Lopez de Prado, and O'Hara (2011) argue that historically high levels of order toxicity forced market makers to withdraw during the Flash Crash. Order flow is considered toxic when it adversely selects market makers who are unaware that they are providing liquidity at their own loss. Easley et al. (2011) develop a metric to measure such toxicity and argue that order flow was becoming increasingly toxic in the hours leading up to the Flash Crash.[58]

There have been a variety of other, smaller illiquidity events in markets since the Flash Crash. On 8June 2011, for example, natural gas futures plummeted 8.1% and then bounced back seconds later. On 2 February 2011, an errant algorithm in oil futures sent 2000-3000 orders in a single second, causing an 8 times spike in volatility and moving the oil price $1 before the algorithm was shut down. In March, trades in 10 new Morningstar ETFs (exchange traded funds) were cancelled when prices fell by as much as 98% following what was determined to be a fat-finger problem (the colloquial name given to an input error entering data).

---

[56] See Madhavan, A., 2011, "Exchange-Traded Funds, Market Structure, and the Flash Crash", Working paper, BlackRock.

[57]See Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun, 2011, The Flash Crash: The Impact of High Frequency Trading on an Electronic Market, SSRN.

[58] See Easley, D., M. Lopez de Prado, and M. O'Hara, 2011, The Microstructure of the Flash Crash: Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading, Journal of Portfolio Management, 37, 118-128.

In summary, the evidence suggests that high frequency trading and algorithmic trading may be contributing to periodic illiquidity in current markets.

5. Future: how is the impact of computer trading on liquidity likely to evolve in the next 10 years?

There is considerable uncertainty regarding the future role of high frequency trading. Tabb Group estimates that HFT accounted for 53% of trading in U.S. markets in the first half of 2011, a decrease from the 61% share it held in 2010. However, the extreme market volatility in August 2011 saw HFT return with a vengeance. Wedbush, one of the largest providers of clearing services to high frequency firms, estimates that high frequency trading made up 75% or more of U.S. volume during the August 4 to 10 period.[59] Whether HFT will continue to be as dominant when volatility subsides to normal levels remains to be seen.

There are also indications that the profitability of HFT is reaching its limits and in the next ten years may come under further pressure.[60] Such reductions may arise for a variety of reasons: the potential move to sub-penny pricing in the U.S. may reduce profitability from market making; new types of multi-venue limit orders may be permitted that will reduce the potential for stale prices across different trading venues; new entrants to the HFT industry will take profits from incumbents; and, regulation and taxation may totally destroy their business model. Lowering costs of entry, which may arise from future technological improvements, can also improve competition. Limiting the value of small improvements in speed by, for example, reducing the value of time priority or requiring a minimum quote life, may also reduce HFT, because it will reduce the incentives for a winner take all speed race

Nonetheless, it seems inevitable that computer trading will remain a dominant force in markets over the next 10 years. One reason for this will be technological advances that facilitate the automated extraction, aggregation, and filtering of news.[61] Such news analytics (NA) could be used in model construction for high frequency traders as well as for portfolio managers. NA technologies currently allow for the electronic “tagging” of news events, corporate filings, and the like, allowing traders with access to this computer technology the ability to see more information faster. Tying such information into computerised trading strategies provides a means for traders to capitalise on information before it is imputed into market prices. HFT will be well positioned to take advantage of such nascent technology.

---

[59] See Mehta, N., “High Frequency Firms Tripled Trades Amid Rout, Wedbush Says”, Bloomberg, August 12, 2011.

[60] See Brogaard (DR10) and Hendershott (DR12).

[61] See Mitra, diBartolomeo, Banerjee, and Yu (DR8).

To the extent such trading puts information into prices more rapidly, markets will benefit by becoming more efficient. However, such strategies also serve to increase the “arms race” in markets by bestowing greater rewards on the most technologically savvy traders. Consequently, it may be that all trading evolves toward computer trading, reflecting that technology diffuses across populations given time.

As this happens, market systems may experience unwanted negative effects.[62] One such effect is already present in the problems of message traffic. Message traffic is the name given to computer instructions to place, change, and cancel orders. On any trading day, message traffic far exceeds trading volume, as far more orders are cancelled or changed than are ever executed. On volatile days, message traffic can cause market outages due to the inability of servers and other computer components of trading to handle the flow. Such outages were widespread during the Flash Crash of May 2010. They recurred in early August 2011, when the extreme volume and volatility took out trading platforms at Goldman Sachs and other large trading firms in the US. When this occurs, market liquidity is affected.

A related systematic risk can arise if a large sub-set of market participants are following the same strategies. For example, if News Analytics (NA) becomes a driving force in portfolio management, then sequences of sell (or buy) orders may arrive at the market, all driven by the same information. For market makers, such orders are “toxic”, because the market maker will be acting as counterparty to agents with better information. As seen in the Flash Crash, when toxicity overwhelms market makers their strategy is to withdraw, and illiquidity results. Consequently, new risk management products will need to evolve to allow market makers, traders, and regulators alike the ability to function.[63] The future of computer trading may thus involve more technology capable of controlling the technology controlling the markets.

6. Conclusions

Computer trading is now the reality in asset markets. Technology has allowed new participants to enter, new trading methods to arise, and even new market structures to evolve. Much of what has transpired in markets is for the good: liquidity has been enhanced, transactions costs have been lowered, and market efficiency appears to be better, or certainly no worse. But there are issues with respect to periodic illiquidity, new forms of manipulation, and potential threats to market stability due to errant algorithms or excessive message traffic that must be addressed. Regulatory changes in practices and policies will be needed to catch up to the new

---

[62] See Farmer and Skouras (DR6).

[63] See, for example, Easley, D., M. Lopez de Prado, and M. O’Hara, 2011, The Exchange of Flow Toxicity, Journal of Trading, 6, 8-13.

realities of trading in asset markets. Caution must be taken to avoid undoing the very many advantages that the high frequency world has brought. Technology will continue to affect asset markets in the future, particularly as it relates to the ultra-fast processing of news into asset prices.

Paper 3: The impact of technology developments

Dave Cliff, University of Bristol.

Key findings

Ongoing advances in the sophistication of ’robot‘ automated trading technology, and reductions in the cost of that technology, are set to continue for the foreseeable future.

Today’s markets involve human traders interacting with large numbers of robot trading systems, yet there is very little scientific understanding of how such markets can behave.

For time-critical aspects of automated trading, readily customisable, special-purpose silicon chips offer major increases in speed; where time is less of an issue, remotely-accessed ’cloud‘ computing services, offer even greater reductions in cost.

Future trading robots will be able to adapt and learn with little human involvement in their design. Far fewer human traders will be needed in the major financial markets of the future.

Executive Summary

The speed and sophistication of computer and communications technology in the financial markets, as elsewhere, is increasing rapidly, while the real cost of the technology is continually falling. There are jobs in the markets that have always been the responsibility of human workers but which can now be done by machines doing the same work for less cost, with fewer errors, and much faster. Moreover, present-day 'robot' computer trading systems are capable of performing jobs that no human trader could ever do, such as assimilating and integrating vast quantities of data and making multiple accurate trading decisions on split-second time-scales. Current technology developments are also providing more sophisticated news analytics techniques. Modern trading systems can learn not only from news, but also from their own experience in the markets. High frequency trading (HFT) is deeply reliant on such technologies.

It is clear that both the pace of development of technology innovations in the financial markets, and the speed of their adoption, look set to continue or increase in the future. Computing power will get cheaper; automated trading systems will get faster, and more intelligent. The availability of cheap 'cloud' computing power means that computers can readily be used to evaluate vast numbers of alternative designs for trading strategies, selecting the best designs and further refining them. The final designs can be implemented not as programmes running on conventional computers, but rather as special-purpose, customised silicon chips, for extra speed. Computer-designed and computer-optimised robot traders are likely to be increasingly viewed as routine, and in time could potentially come to replace current algorithms designed and refined by humans. Future cloud computing services could provide automated compilation down to customised silicon-chip hardware.

Already, a very large proportion of transactions in the markets are computer-generated, yet large numbers of human traders remain in the markets. In several significant markets (such as foreign exchange) the overall number of human traders engaged in on-the-spot execution of orders has fallen sharply in recent years, and is likely to continue to reduce in the future. Nevertheless, the present-day mix of human and robot traders looks set to continue for some time. Given this mix of humans and robots, studying present-day markets lies neither wholly within the realms of social sciences such as economics, nor wholly within the realm of computer and communications systems engineering. Rather, the current markets are manifestly *socio-technical systems*. The characteristics and dynamics of markets populated by mixtures of humans and robots are not at all well understood, and they sometimes behave in unpredictable, undesirable ways. Even fully automated trading systems, with 'robot traders' provided by many different parties, need very careful study. The primary challenges for the future are centred on how the complex dynamical socio-technical ecosystem of the global financial markets can be mapped, managed, and modified to prevent undesirable behaviour[1].

These technology developments mean that major trading systems can today exist anywhere. Emerging economies may capitalise on the opportunities made available by new technologies and thereby may come to threaten the long-established historical dominance of major European and US cities as global hubs for trading in the financial markets.

1: Introduction

The present-day move to ever higher degrees of automation on the trading floors of exchanges, banks, and fund-management companies is similar to the major shift to automated production and assembly that manufacturing engineering underwent in advanced economies during the 1980s and 1990s, and is likely to have a similar effect on the distribution of employment in the financial sector. Already, a very large proportion of transactions in the markets are computer-generated and yet the characteristics and dynamics of markets populated by mixtures of human traders and machine traders are not at all well understood, and sometimes the markets behave in unpredictable, undesirable ways. Few details are known of the connectivity network of interactions and dependencies in technology-enabled financial markets. There is a clear need to map the current global financial network, to gain an understanding of the current situation. Such a mapping will enable the development of new tools and techniques for managing the financial network, and to explore how it can be modified to reduce or prevent undesirable behaviour.[64] It is clear that new technology, new science and engineering tools and techniques, will be required to help map, manage, and modify the market systems of the future.

2: How has financial market technology evolved?

The technology changes of the past five years are best understood as a continuation of longer-term trends. Cliff, Brown, and Treleaven (DR3) relate the history of technology in the financial markets, covering the 18th, 19th, and 20th centuries in brief and then explore in more detail the rapid and significant changes which have occurred in the opening years of the 21st century. Their narrative is summarised here, which draws on many sources.[65]

---

[64] The need to map, manage, and modify the financial network is the central message from a 2009 speech *Rethinking the Financial Network* by A. Haldane, the Bank of England's Executive Director responsible for Financial Stability: see http://www.bankofengland.co.uk/publications/speeches/2009/speech386.pdf.

[65] For other accounts of the recent history of technology developments in the financial markets, the following three texts are particularly recommended: J. Angel, L. Harris, & C. Spratt (2010), *Trading in the 21st Century*. (Unpublished manuscript, available from http://www.sec.gov/comments/s7-02-10/s70210-54.pdf); P. Gomber, *et al.* (2011). *High Frequency Trading*. Technical Report, Goethe Universität & Deutsche Börse; and D. Leinweber (2009), *Nerds on Wall Street*. John Wiley Publishers. For a very recent discussion of high-frequency trading, including interviews with leading practitioners, see E. Perez (2011), *The Speed Traders*. McGraw-Hill Publishers.

As Cliff *et al.* discuss in DR3, the high-speed processing of data, and high-speed communication of data from one location to another, have always been significant priorities for the financial markets**.** Long before the invention of computers or pocket-calculators, traders with fast mental arithmetic skills could out-smart their slower-witted competitors. In the 19th Century, communication of financially significant information by horse-riding messengers was replaced by the faster 'technology' of carrier pigeons; then pigeons were made redundant by telegraph; and then telegraph by telephones. In the last quarter of the 20th century, the shift to computer-based trading systems meant that automated trading systems could start to perform functions previously done only by humans: computers could monitor the price of a financial instrument (a share-price, say) and could issue orders to buy or sell the instrument if its price rose above or below specified 'trigger' prices. Such very simple 'program trading' systems were widely blamed for the Black Monday crash of October 1987, the memory of which for several years afterwards dampened enthusiasm for allowing computers to issue buy or sell orders in the markets. Nevertheless, the real cost of computer power continued to halve roughly once every two years (the so-called 'Moore's Law' effect), and so by the late 1990s it was possible to buy, at no extra cost, computers roughly 100 times more powerful than those used in 1987. All this extra computer-power could be put to use in implementing much more sophisticated processing for making investment decisions, and for issuing structured patterns of orders to the markets.

By the turn of the millennium, as the real cost of computing continued to fall at dramatic pace, the management of investment funds had become an increasingly technical field, heavily dependent on computationally intensive mathematical models to reduce or offset portfolio risk, i.e. to 'hedge' the risk in the fund's holdings, giving rise to the phrase 'hedge fund'. Many hedge funds based their investment decisions on so-called *statistical arbitrage* (commonly abbreviated to 'stat arb'). One popular class of stat arb strategies identify long-term statistical relationships between different financial instruments, and trade on the assumption that any deviations from those long-term relationships are temporary aberrations, that the relationship would revert to its mean in due course. One of the simplest such 'mean-reversion' strategies is *pairs trading*, where the statistical relationship which is used as a trading signal is the degree of correlation between just two securities. Identifying productive pair-wise correlations in the sea of financial-market data is a computationally demanding task, but as the price of computer-power fell, so it became possible to attempt ever more sophisticated stat arb strategies.

At much the same time, the availability of cheaper computation meant that it was possible to deploy automated trading systems that had considerably more intelligence than those implicated in the 1987 crash. In most cases, this intelligence was based on rigorous mathematical approaches that were firmly grounded in statistical modelling and probability theory. The new wave of automated systems concentrated on *execution* of a trade. The computer did not make the decision to buy or to sell a particular block of shares or quantity of commodity, nor to convert a particular amount of one currency into another: those decisions were still taken by humans (possibly on the basis of complex statistical analysis). But, once the trading decision had been made, the execution of that trade was then handed over to an automated execution system (AES). Initially, the motivation for passing trades to an AES was that the human traders were then freed up for dealing with more complicated trades. As AES became more commonplace, and more trusted, various trading institutions started to experiment with more sophisticated approaches to automated execution: different methods, different *algorithms*, could be deployed to fit the constraints of different classes of transaction, under differing market circumstances; and hence the notion of 'algorithmic trading' was born.

At the same time as AES systems were being developed to reduce market impact, other trading teams were perfecting advanced stat arb techniques for identifying trading opportunities based on complex statistical regularities which lay deep in the data: the price and volume data for hundreds or thousands of instruments might have to be considered simultaneously and cross-compared, in the search for opportunities similar to the pairs trading of the 1980's, but typically involving very many more than two instruments. These advanced stat arb approaches were made possible by powerful computers used to run the statistical analyses, and also by developments in computer-based trading infrastructure (the machinery which traders use to communicate with each other and with the exchanges). Two notable developments were *Straight-Through Processing* (STP), where the entire trading process from initiation of an order to final payments and clearing is one seamless electronic flow of transaction-processing steps with no human-operated intermediate stages; and *Direct Market Access* (DMA), where investors and investment funds are given direct access to the electronic order-books of an exchange, rather than having to interact with the market via an intermediary such as an investment bank or broker/dealer.

The convergence of cheap computer-power, statistically sophisticated and computationally intensive trading strategies, fast automated execution via STP, and DMA, means that in the last two or three years it has become commonplace for market participants to seek counterparties to a transaction electronically, identify a counterparty, and then execute the transaction, all within a small number of seconds.

The old "vertically integrated" business model of investment banking is becoming increasingly fragmented. One effect of the EU's MiFID legislation was to create an ecosystem of small and medium-sized businesses offering 'middleware' technology components that could each be purchased and then plugged together to achieve the same functionality which had previously been the exclusive preserve of the trading systems developed in-house by big institutions. This lowered the barriers to entry: armed with enough cash, one or two entrepreneurs working in a rented office with a high-speed internet connection can set up a trading company and automate much, or perhaps all, of the workflow required to run a fund. At the same time, a new style of trading has emerged, known as high-frequency trading (HFT),[66] where automated systems buy and sell on electronic exchange venues, sometimes holding a particular position for only a few seconds or less. That is, an HFT system might ‘go long’ by buying a quantity of shares (or some other financial instrument, such as a commodity or a currency) hold it for perhaps two or three seconds, and then sell it on to a buyer. If the price of the instrument rises in those two or three seconds, and so long as the transaction costs are small enough, then the HFT system has made a profit on the sale. The profit from holding a long position for three seconds is unlikely to be great, and it may only be a couple of pennies, but if the HFT system is entirely automated, then it is a machine that can create a steady stream of pennies per second, of dollars per hour, twenty four hours per day. A recent study by Kearns *et al.* (2011)[67] indicates that the total amount of money extractable from the markets via HFT may be more modest than some might estimate or guess: a few tens of billions of dollars in the US markets. Despite this, the low variation in positive returns (the ‘steady’ in ‘steady stream of pennies’) from a well-tuned HFT system is an attractive feature, and one that makes HFT an area of intense interest in the current markets.

As the global financial markets became dependent on computers running automated trading systems and communicating with each other over optical fibre networks, the speeds of computation and of communication became two primary means by which competitive advantage could be gained and maintained. The effect of this in the present-day markets is discussed in the next section.

3: What are the key current technology developments?

Firms at the front line of the financial markets, such as investment banks, fund management companies, and exchange operators, are all critically dependent on information technology (IT) and the telecommunications networks that allows computers to talk to each other. For the past two decades, nearly all such firms used their own in-house IT systems, very often involving powerful ’server‘ computers connected to ’client‘ computers running on the desks of each employee. Almost always, the client computers would be standard personal computers (PCs), of the sort available from high-street retailers; and the server computers would actually be constructed from several very high-specification PC computers, all located together and

[66] Google Trends (trends.google.com) indicates that Google’s users, worldwide, have only used the phrase “high frequency trading” as a search term in the past three years, and “algorithmic trading” only in the past five years.

[67] M. Kearns, A. Kulesza, & Y. Nevmyvaka (2010). Empirical Limitations on High Frequency Trading Profitability. *The Journal of Trading,* **5**(4):50-62. Available from http://www.cis.upenn.edu/~mkearns/papers/hft_arxiv.pdf.

connected to each other in a single room; that room being the firm’s ’server room‘, or ’data centre‘.

As Cliff, Brown, and Trelevean (DR3) describe in some detail, the global IT industry is currently undergoing a major shift, toward ‘cloud computing‘ where ultra-large-scale data-centres (truly vast warehouses full of interconnected computers) are accessed remotely as a service via the internet, with the user of the remotely-accessed computers paying rental costs by the minute or by the hour. This greatly lowers the cost of high-performance computing (HPC), and hence reduces barriers to entry for individuals or firms looking to use supercomputer-scale HPC for the automated design and optimisation of trading systems: rather than spend millions of dollars of capital expenditure on an in-house HPC data-centre facility, it is now possible to get the same results from a few thousand dollars of renting HPC from cloud-computing providers. That is, it is no longer necessary to have the financial resources of a major hedge fund or investment bank to engage in development of highly technology-dependent approaches to trading. The full implications of this are not yet clear.

At the same time, the desire for ultra-high-speed processing of financial data has led a number of market leaders to abandon the use of general-purpose computers such as commercially available PCs, and replace them instead with customised special-purpose silicon chips. Some of these silicon chips have hundreds or thousands of independent small computers on them, each operating in parallel, giving huge increases in speed. This is discussed in more depth by Cliff *et al.* (DR3), which includes discussion of financial companies which have already made this move, concluding that the switch to such ‘custom silicon’ is set to continue in the coming decade.

These two trends mean that greater computational power and greater speed are becoming more readily available per unit cost, and so technologists have turned their attention to developing innovative new systems for automated generation of trading decisions and/or automated execution of the orders necessary to enact those decisions.

One major new technology that is currently the focus of significant research and development is the prospect of computers being programmed to ‘understand’ not only the numeric information of market prices, volumes, and times, but also the non-numeric ‘semantic’ information that is carried in human-readable data-streams such as written news reports; audio data such as telephone calls, radio shows, podcasts,and video sequences. This is an issue explored in depth by Mitra *et al.,* whose work is summarised in Section 3.1.

Despite the increases in computer power, processing speed, and sophistication of computer algorithms, the present-day financial markets still involve large numbers of human traders. There are good reasons for expecting that for the next decade or so the number of humans participants in the market will remain significant. For most major markets in the US, UK, and mainland Europe, the proportion of computer-generated trades is estimated to be variously 30%, 50%, and in some cases nearer 75%[68]. What is clear then, is that the current markets involve large numbers of human traders interacting with large numbers of automated trading systems. This is a major shift in the make-up of the markets, and may well have affected the dynamics of the markets, (see Chapters 1 and 2). Research that explores the interactions of humans and algorithmic trading systems is discussed in Section 3.2.

3.1 Automated analysis of market news and sentiment

Mitra *et al.* (DR8) commence the core of their report with a consideration of which asset classes are best suited to automated trading by computers. They contend that different financial instruments have different liquidity and that the optimal trading frequency for an instrument can be expressed as a function of the liquidity of its market, amongst other factors. The higher the optimal trading frequency, the more useful algorithmic trading is. The traders in a market can be classified as one of two types: those who aim to profit by merely providing liquidity (so-called 'inventory traders'); and those who instead aim to profit by trading on the basis of information. Inventory traders act as 'market-makers': they hold a sufficiently large quantity of an instrument (their inventory) that they are always able to service buy or sell requests, and they make money by setting a higher price for selling than for buying (this is the type of business model that is familiar from any airport foreign-exchange retail outlet). Inventory traders can, in principle, operate profitably without recourse to any information external to the market in which their instruments are being traded. The second class of traders, known as 'informed' or 'value-motivated' traders, make use of information in news stories and related discussion and analysis to come to a view about what price an instrument should be trading at either now or in the future, and then buy or sell that instrument if their personal opinion on the price is different from the current market value. In recent years, technologies have been developed that allow computers to analyse news stories and discussions on social networking websites, and they are rapidly increasing in sophistication.

Mitra *et al.* argue that the primary asset classes suitable for automated trading are equities, including exchange-traded funds (ETFs) and index futures, foreign exchange, and to a lesser extent commodities, and fixed income instruments. ETFs are securities traded on major exchanges as if they were standard stock equities (shares in a company), but the ETF instead represents a share in a holding of assets such as commodities, currency, or stock. As would be expected, news events (both anticipated and unanticipated) can affect both traditional manual trading for these asset classes and also automated trading activities. Anticipated events are those such as the release of official inflation data by government treasury

---

[68] See, e.g.: I. Kaminska (2011), Algo trading and the Nymex, *Financial Times Alphaville Blog,* http://ftalphaville.ft.com/blog/2011/03/04/505021/algo-trading-and-the-nymex/, and I. Kaminska (2009), HFT in Europe, *Financial Times Alphaville Blog*, http://ftalphaville.ft.com/blog/2009/07/24/63651/high-frequency-trading-in-europe/.

departments or scheduled earnings announcements by firms; unanticipated events are those such as news concerning major accidents, terrorist actions, or Act-of-God natural disasters. Because of the effect that news events can have on the prices of financial instruments, major global companies exist to provide news feeds specific to the financial markets, including Thompson Reuters, The Financial Times, The Wall Street Journal, Dow Jones Newswire and Bloomberg. Much of the news content comes in formats that can readily be processed by computers. The content from traditional mass-market news broadcasters, such as the BBC, can also be processed by computers (possibly after some automated re-formatting, or conversion from audio/video to text-based transcripts).

Because of this, researchers in academia and in financial institutions have developed methods for news analytics. Significant advances have been made in recent years, and the techniques are still growing in sophistication. In general, it is reasonable to predict that a computer will be able to react to a breaking news story faster than a human can, but of course this is only useful if its analysis of the story is actually correct. Some practitioners argue that automated trading and news analytics puts manual (human-based) trading at a considerable disadvantage; and that this applies to both retail investors and institutional investors. Although the current state of the art in leading-edge news analytics technology can not yet reliably outperform a well-informed human trader reading the same material, and has only very limited abilities in comparison to the human capacity for reasoning and lateral thinking, the capabilities and sophistication of news analytics systems will continue to increase over the next decade, possibly to the point where they surpass the performance of human analysts and traders.

3.2 Studying interactions between human and algorithmic trading systems

The global financial markets are now populated by two types of economic agent: human traders, and ‘software agents’. The latter are either algorithmic systems performing trading jobs that ten or 20 years ago would have been the responsibility of humans, or HFT systems doing jobs that no human could ever hope to attempt. Interactions between human traders in electronic markets has long been studied in the field known as *Experimental Economics,* and more recently the interactions between software-agent traders in electronic markets has been the topic of various research studies in so-called *Agent-based Computational Economics*. Curiously, these two research fields are largely distinct: the first studies markets populated entirely by human traders while the second studies markets populated entirely by algorithmic software-agent traders. There is a surprisingly stark lack of studies of the interactions between human traders and algorithmic trading systems. That is, there is really very little scientific literature that explores heterogeneous markets, populated by both humans and algorithmic systems. De Luca *et al.* (DR 13) surveyed the surprisingly small amount of published peer-reviewed literature which does describe scientific studies of interactions between human and algorithmic traders.

The first paper to report any such results was published in 2001[69] by a team of researchers working at IBM, where two trading algorithms were each demonstrated to outperform human traders. IBM's work served as an inspiration to Prof. Jens Grossklags of the University of California at Berkeley, who used similar methods to explore different questions in papers published in 2003 and 2006[70]. Until 2011, the experiments reported by IBM and by Grossklags were the only three peer-reviewed papers in the scientific literature that studied this topic. This is a startling 'gap' in the research literature which has only very recently started to be filled. The most recent papers in this field are two published by De Luca and Cliff in 2011[71], replicating and extending IBM's experiments from 2001.

In DR 13, De Luca *et al. c*ontend that the relative lack of such studies is a serious omission from the literature. They give a detailed description and analysis of results from several new experiments, conducted specifically for the Foresight driver review (DR13), where some of the artificial experimental constraints that were used in earlier work are relaxed, for greater realism and hence increased relevance to the real-world markets. The key conclusion from De Luca *et al.*'s review is that their new experiments indicate that the previously reported outperformance of the algorithmic trading systems over humans are probably a consequence of the artificial nature of the experimental designs in which they were evaluated, and may also be due simply to the fact that computers can act and react faster than humans. When the flow of orders in the market was trickled in gradually (rather than all orders being released simultaneously, which was an artificial constraint in the designs of the earlier experiments) the performance of the software agents relative to the humans was significantly diminished. De Luca *et al.*'s results in DR13 provide some empirical support for the intuitive notion that the primary advantage that current software-agent trading algorithms have over humans is the speed at which they operate, although being faster at trading does not necessarily lead to greater overall efficiency in the market.

3.3 From the present to the future

Some key issues in today's markets look likely to continue to remain vitally important in the future. For instance, cyber-security will remain a core concern: electronic attacks on the computer systems and communications

---

[69] R Das, J. Hanson, J. Kephart, & G. Tesauro (2001). Agent-human interactions in the continuous double auction. In *Proceedings of IJCAI-01*,
http://www.research.ibm.com/infoecon/paps/AgentHuman.pdf.

[70] J. Grossklags & C. Schmidt (2006). Software agents and market (in)efficiency: a human trader experiment. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews),* **36**(1):56-67.

[71] See: M. De Luca & D. Cliff (2011a). Agent-Human Interactions in the Continuous Double Auction, Redux: Using the OpEx Lab-in-a-Box to explore ZIP and GDX. In *Proceedings of the Third International Conference on Agents and Artificial Intelligence (ICAART-2011);* and M. De Luca & D. Cliff (2011b). Human-Agent Auction Interactions: Adaptive-Aggressive Agents Dominate. *Proceedings of IJCAI-2011.*

networks of the global financial markets are always likely to be attractive to criminals. Furthermore, the widespread move toward higher reliance on advanced computing technology means that the speed of light is now a significant limiting factor in determining how trading systems interact with one another. Even with the best technology imaginable, information cannot be transmitted faster than the speed of light, and even at light-speed it does actually take measurable periods of time to move information across an ocean, or even across a few city blocks. This constraint is never going to go away.

The evolution of classes of algorithms is discussed in detail by Gergely Gyurkó, in DR5. Recent academic research indicates that coming generations of trading algorithms will be adaptive (learning from their own experience) and will be the result of automated computerised design and optimisation processes. Therefore, the performance of next-generation trading algorithms may be extremely difficult to understand or explain, both at the level of an individual algorithm, and at the system-level of the market itself.

4: Technology advances likely in the next 10 years

As with many such competitive interactions, in the technology arms-race new innovations only confer competitive advantage to an innovator for as long as it takes the innovator's competitors to copy that innovation or to come up with counter-innovations of their own: as soon as all traders are using a particular new technology, the playing field is levelled again. Nevertheless, several of the present-day technology trends seem likely to remain significant factors over the coming years.

4.1 Cloud computing

Cloud computing offers the possibility that it is no longer necessary to have the financial resources of a major hedge fund or investment bank to engage development of highly technology-dependent approaches to trading. Nevertheless, there are regulatory and legislative issues that need to be carefully examined: for example for jurisdictional reasons, the geographic location of the remote servers can matter greatly. Cloud-computing service providers are well aware of such concerns, and can offer geographic guarantees in their service-level agreements and contracts. Moreover, remote access of computing facilities, even at the speed of light, means that there will be latencies in accessing the remote systems. For very many applications, these may not matter, but for trading activities, the latencies inherent in communicating with remote data-centres can be prohibitive. Latency would certainly be a problem if an institution tried to run its automated HFT algorithms 'in the cloud', but it is important to remember that not all trading is HFT: there are other modes of trading, such as long-only macro trading, that are not so latency-sensitive.

The primary impact of cloud computing on activities in the financial markets in the next ten years will not be in the provision of computing facilities that automate execution, but rather in the ability of the cloud to provide cheap, elastically scalable, high-performance computing (HPC). Such cheap, remote HPC will allow massively computer-intensive procedures to be deployed for the automated design and optimisation of trading strategies and execution algorithms: this computational process is not latency sensitive. Many major investment banks and hedge funds already own and operate their private data-centres, but they do this for business-critical operations and only a fraction of the capacity can be turned to HPC uses. The ability to either extend existing in-house computer power by adding on cloud-based resources (known as 'cloudbursting') or to simply outsource all of the HPC provisioning to a cloud provider, opens up new possibilities that are only just being explored.

4.2 Custom silicon

General-purpose, commercially available personal computers (PCs) have over recent years moved from being based on a single central processing unit (CPU) chip such as an Intel Pentium, to a new breed of CPU chips that have multiple independent computers (known as "cores") built into them. Perusal of high-street stores will reveal PCs with dual-core and with quad-core chips as standard. In rough terms, a dual-core chip can do twice as much work as a single-core chip per unit of time, and a quad-core can do four times as much. Currently there is a major shift underway toward so-called many-core computing, exploring the speed-ups offered by using chips with many tens or hundreds of independent cores operating in parallel: often this involves using specialised processing chips originally designed for computer graphics processing. Furthermore, as was noted above, the desire for ultra-high-speed processing of financial data has led a number of market leaders to abandon the use of general-purpose computers such as commercially available PCs, and replace them with special-purpose silicon chips that can be customised or programmed 'in the field' (i.e., the end-user of the chip customises it to whatever purpose fits that user's needs). Currently the most popular such technology is a type of chip known as a field-programmable gate array (FPGA). Currently, programming an FPGA is a very complicated and time-consuming task: the programmer has to translate an algorithm into the design for an electronic circuit and describe that design in a specialised hardware description language. Despite these complexities, the switch to such custom silicon is likely to continue over the next decade because of the speed gains that it offers. Over that period it is probable that the use of FPGAs will be supplanted by a newer approach to custom silicon production, involving more readily field-programmable multi-core or many-core chips. Such chips will be programmable in a high-level software language much like current industry-standard programming languages. This means that conventionally-trained programmers can write algorithms that are then 'compiled down' onto the underlying silicon-chip hardware, without the need to learn specialised FPGA hardware description languages. This has the potential to reduce custom-silicon development times (currently measured in days or weeks) down to waits of only a few minutes from describing a trading algorithm in a high-level programming language, to having that algorithm running on a massively parallel high-speed computing array composed of many independent customised silicon-chip processing elements. In DR3, Cliff, Brown, and Treleaven claim that this style of computer hardware is likely to be in wide use within the next decade. This type of hardware would have enough computing power to enable future generations of trading algorithms that are adaptive (learning from their own experience) and that will not have been designed by human engineers, but rather will be the result of automated computerised design and optimisation processes.

4.3 Computer-generated trading algorithms that adapt and learn

The use of automated optimisation methods to design and improve autonomous adaptive trading algorithms is already commonplace in academic research, and its use in the finance industry looks set to increase over the next decade. This is a development that is enabled and accelerated by the step-change drop in cost of HPC offered by cloud computing service providers, and by the huge speed increases offered by custom silicon. Because these next-generation trading algorithms will have little or no human involvement in their design and refinement, the behaviour of any one such automated trader may be extremely difficult to understand or explain, and the dynamics of markets populated by such traders could be very difficult to predict or control.

5: Conclusion

It is reasonable to speculate that the number of human traders involved in the financial markets could fall dramatically over the next ten years. While unlikely, it is not impossible that human traders will simply no longer be required at all in some market roles. The simple fact is that we humans are made from hardware that is just too bandwidth-limited, and too slow, to compete with coming waves of computer technology.

Just as real physical robots revolutionised manufacturing engineering, most notably in automobile production, in the latter years of the 20th Century, leading to major reductions in the number of employees required at car plants, so the early years of the 21st seem likely to be a period in which a similar revolution (involving software robot traders) occurs in the global financial markets. The number of front-line traders employed by major financial institutions is likely to fall, but there may be increased demand for developers of algorithms.

On the basis of the evidence reviewed in the various papers discussed in this chapter, it is clear that both the pace of development of technology innovations in the financial markets, and the speed of their adoption, look set to continue or increase in the future. One stark implication of the developments reviewed here is highlighted in DR3: trading systems can today exist anywhere. Emerging economies such as those of Brazil, Russia, India, and China may capitalise on the opportunities available from the new technologies and in doing so may, within only a few decades, come to threaten the historical dominance of major European cities as global hubs for financial trading. Formulating appropriate policy responses to such potential threats is a matter for further consideration.

Glossary of terms

Volatility – variability of an asset's price over time, often measured in percentage terms.

Financial market stability – the lack of extreme movements in asset prices over short time periods.

Liquidity – the ability to buy or sell an asset without greatly affecting its price. The more liquid the market, the less the price impact will be.

Market making – providing liquidity to buyers and sellers by acting as a counterparty. A market maker buys from sellers and sells to buyers.

Designated liquidity provider – a general term given to a market participant who agrees to stand ready to buy or sell an asset to accommodate market demand.

Order book – the collected limit orders to buy or sell an asset. Order books today are generally electronic and allow traders to specify a price at which they would like to buy a specified quantity of an asset or a price at which they would like to sell a quantity of the asset.

Order flows – the arrival of buy orders and sell orders to the market

Market efficiency – concept that market prices reflect the true underlying value of the asset.

Transaction costs – the costs trader incur to buy or sell an asset

Market transparency – the ability to see market information. Post-trade transparency refers to the ability to see trade prices and quantities. Pre-trade transparency refers to the ability to see quotes.

Price efficiency – when an asset's price reflects the true underlying value of an asset.

Price discovery – the market process whereby new information is impounded into asset prices.

Foresight
1 Victoria Street
London SW1H 0ET
www.bis.gov.uk
URN 11/1276