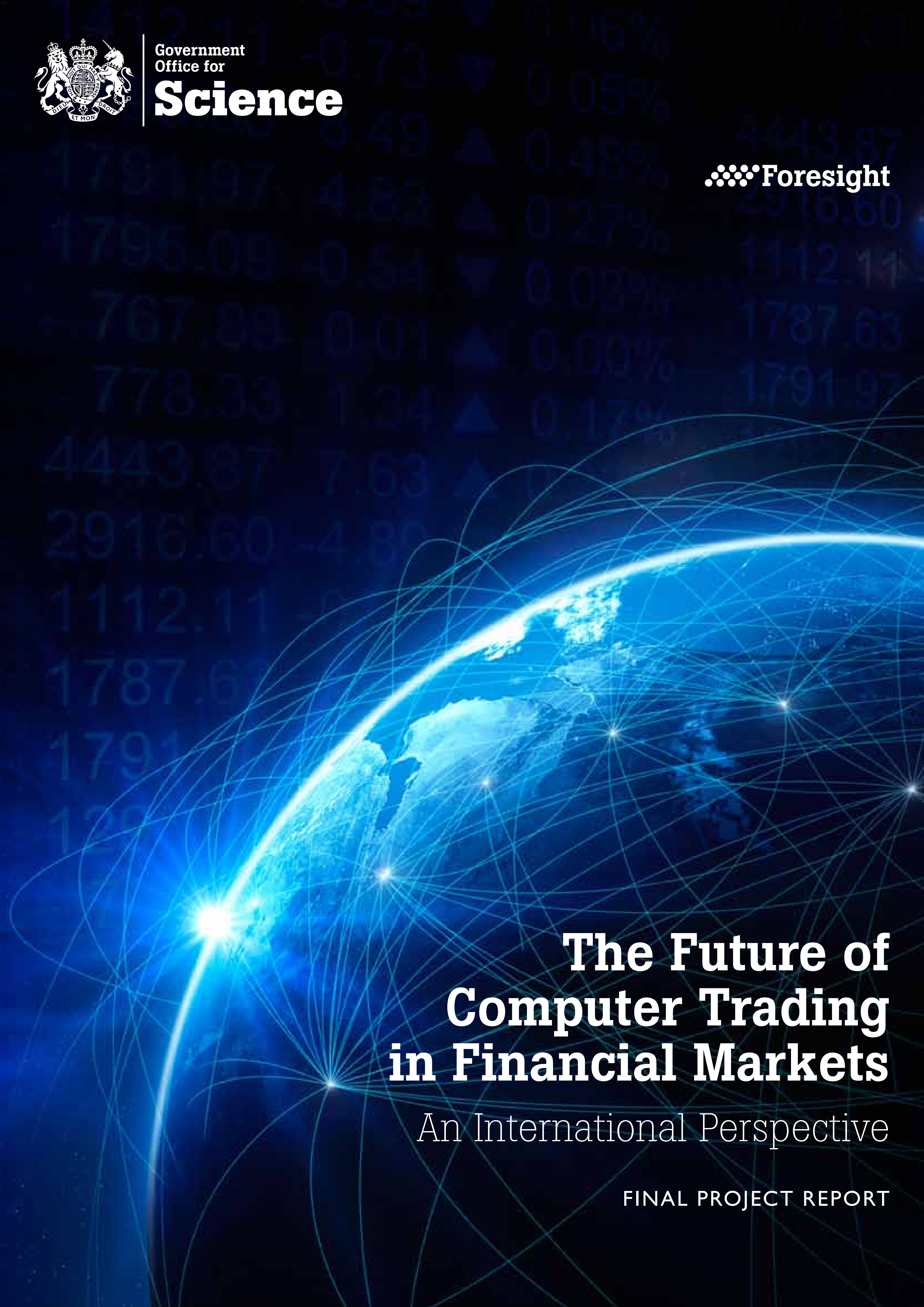
Government
Office for
Science

Foresight



The Future of Computer Trading in Financial Markets

An International Perspective

FINAL PROJECT REPORT

This Report should be cited as:

Foresight: The Future of Computer Trading in Financial Markets (2012)
Final Project Report
The Government Office for Science, London

The Future of Computer Trading in Financial Markets

An International Perspective

This Report is intended for:

Policy makers, legislators, regulators and a wide range of professionals and researchers whose interest relate to computer trading within financial markets. This Report focuses on computer trading from an international perspective, and is not limited to one particular market.

Foreword

Well functioning financial markets are vital for everyone. They support businesses and growth across the world. They provide important services for investors, from large pension funds to the smallest investors. And they can even affect the long-term security of entire countries.

Financial markets are evolving ever faster through interacting forces such as globalisation, changes in geopolitics, competition, evolving regulation and demographic shifts. However, the development of new technology is arguably driving the fastest changes. Technological developments are undoubtedly fuelling many new products and services, and are contributing to the dynamism of financial markets. In particular, high frequency computer-based trading (HFT) has grown in recent years to represent about 30% of equity trading in the UK and possible over 60% in the USA.

HFT has many proponents. Its roll-out is contributing to fundamental shifts in market structures being seen across the world and, in turn, these are significantly affecting the fortunes of many market participants. But the relentless rise of HFT and algorithmic trading (AT) has also attracted considerable controversy and opposition. Some question the added value it brings to markets and, indeed, whether it constitutes a drag on market efficiency. Crucially, some also believe that it may be playing an increasing role in driving instabilities in particular markets. This is of concern to all financial markets, irrespective of their use of HFT, since increasing globalisation means that such instabilities could potentially spread through contagion. It has also been suggested that HFT may have significant negative implications relating to market abuse. For these reasons, it is unsurprising that HFT is now attracting the urgent attention of policy makers and regulators across the world.

This international Foresight Project was commissioned to address two critical challenges. First, the pace of technological change, coupled with the ever-increasing complexity of financial trading and markets, makes it difficult to fully understand the present effect of HFT/AT on financial markets, let alone to develop policies and regulatory interventions that are robust to developments over the next decade. Second, there is a relative paucity of evidence and analysis to inform new regulations, not least because of the time lag between rapid technological developments and research into their effects. This latter point is of particular concern, since good regulation clearly needs to be founded on good evidence and sound analysis.

Therefore, the key aim of this Project has been to assemble and analyse the available evidence concerning the effect of HFT on financial markets. Looking back through recent years and out to 2022, it has taken an independent scientific view. The intention has been to provide advice to policy makers. Over 150 leading academics from more than 20 countries have been involved in the work which has been informed by over 50 commissioned papers, which have been subject to independent peer review.

The key message is mixed. The Project has found that some of the commonly held negative perceptions surrounding HFT are not supported by the available evidence and, indeed, that HFT may have modestly improved the functioning of markets in some respects. However, it is believed that policy makers are justified in being concerned about the possible effects of HFT on instability in financial markets. Therefore, this Report provides clear advice on what regulatory measures might be most effective in addressing those concerns in the shorter term, while preserving any benefits that HFT/AT may bring. It also advises what further actions should be undertaken to inform policies in the longer term, particularly in view of outstanding uncertainties. In conclusion, it is my pleasure to make this Report and all of its supporting evidence and analysis freely available. It is my hope that it will provide valuable insights into this crucial issue.

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 (Chair)	Non-executive Director, Legal & General plc, Amadeus IT Holding SA, Nomura Holdings Inc., Chairman, Nomura Bank International, Non-executive Director, Department for Work and Pensions and Senior Adviser, Chatham House.
Professor Philip Bond	Visiting Professor of Engineering Mathematics and Computer Science at the University of Bristol and Visiting Fellow at the Oxford Centre for Industrial and Applied Mathematics.
Professor Dave Cliff	Professor of Computer Science at the University of Bristol.
Professor Charles Goodhart CBE, FBA	Professor (Emeritus) of Banking and Finance at the London School of Economics.
Kevin Houstoun	Chairman of Rapid Addition and co-Chair of the Global Technical Committee, FIX Protocol Limited.
Professor Oliver Linton FBA	Chair of Political Economy at the University of Cambridge.
Dr Jean-Pierre Zigrand	Reader in Finance at the London School of Economics.

Foresight would like to thank Dr Sylvain Friederich, University of Bristol, Professor Maureen O'Hara, Cornell University and Professor Richard Payne, Cass Business School, City University, London for their involvement in drafting parts of this Report.

Foresight would also like to thank Andy Haldane, Executive Director for Financial Stability at the Bank of England, for his contribution in the early stages of the Project.

Foresight Project team:

Professor Sandy	Thomas	Head of Foresight
Derek	Flynn	Deputy Head of Foresight
Lucas	Pedace	Project Leader
Alexander	Burgerman	Project Manager
Gary	Cook	Project Manager
Christopher	Griffin	Project Manager
Anne	Hollowday	Project Manager
Jorge	Lazaro	Project Manager
Luke	Ryder	Project Manager
Piers	Davenport	Project Co-ordinator
Martin	Ford	Project Co-ordinator
Yasmin	Hossain	Project Researcher
Zubin	Siganporia	Project Researcher
Isabel	Hacche	Intern
Arun	Karnad	Intern
Louise	Pakseresht	Intern
Jennifer	Towers	Intern

For further information about the Project please visit:

http://www.bis.gov.uk/foresight

Contents

Executive summary

A key message: despite commonly held negative perceptions, the available evidence indicates that high frequency trading (HFT) and algorithmic trading (AT) may have several beneficial effects on markets. However, HFT/AT may cause instabilities in financial markets in specific circumstances. This Project has shown that carefully chosen regulatory measures can help to address concerns in the shorter term. However, further work is needed to inform policies in the longer term, particularly in view of likely uncertainties and lack of data. This will be vital to support evidence-based regulation in this controversial and rapidly evolving field.

1 The aims and ambitions of the Project

The Project's two aims are:

- to determine how computer-based trading (CBT) in financial markets across the world could evolve over the next ten years, identifying potential risks and opportunities that this could present, notably in terms of financial stability[1] but also in terms of other market outcomes, such as volatility, liquidity, price efficiency and price discovery;
- to draw upon the available science and other evidence to provide advice to policy makers, regulators and legislators on the options for addressing present and future risks while realising potential benefits.

An independent analysis and an international academic perspective:

The analysis provides an independent view and is based upon the latest science and evidence. As such, it does not constitute the views or policy of the UK or any other government.

Over 150 leading academics and experts from more than 20 countries have been involved in the work which has been informed by over 50 commissioned scientific papers, which have been independently peer reviewed. A further 350 stakeholders from across the world also provided advice on the key issues to consider[2].

2 Why the Project was undertaken

Well functioning financial markets are vital for the growth of economies, the prosperity and well-being of individuals, and can even affect the security of entire countries. Markets are evolving rapidly in a difficult environment, characterised by converging and interacting macro- and microeconomic forces, such as globalisation, changes in geopolitics, competition, evolving regulation and demographic shifts. However, the development and application of new technology is arguably causing the most rapid changes in financial markets. In particular, HFT and AT in financial markets have attracted considerable controversy relating to their possible benefits and risks.

While HFT and AT have many proponents, others question the added value they bring to markets, and indeed whether they constitute a drag on market efficiency. Crucially, some believe they may be playing an increasingly significant role in driving instabilities in particular markets. There have been suggestions that HFT and AT may have significant negative implications relating to market abuse. For these reasons, and in view of the vital importance of financial markets, both HFT and AT are now attracting the urgent attention of policy makers and regulators across the world.

1 A list of definitions used in this Executive Summary can be found in Annex C of the Project's Final Report.

2 A list of individuals who have been involved can be found in Annex A of the Project's Final Report.

Two challenges for regulators:

Effective regulation must be founded on robust evidence and sound analysis. However, this Project addresses two particular challenges currently faced by regulators:

- *Rapid developments and applications of new technology, coupled with ever-increasing complexity of financial trading and markets make it difficult to fully understand the present effects of HFT and AT on financial markets and even more difficult to develop policies and regulatory interventions which will be robust to developments over the next decade.*
- *There is a relative lack of evidence and analysis to inform the development of new regulations, not least because of the time lag between rapid technological developments and research into their effects, and the lack of available, comprehensive and consistent data.*

These two challenges raise important concerns about the level of resources available to regulators in addressing present and future issues. Setting the right level of resources is a matter for politicians. However, unlocking the skills and resources of the wider international academic community could also help. Here, a drive towards making better data available for analysis should be a key objective and the experience of this Project suggests that political impetus could be important in achieving that quickly.

It makes sense for the various parties involved in financial markets to be brought together in framing further analytical work, in order to promote wide agreement to the eventual results. Everyone will benefit from further research that addresses areas of controversy, as these can cloud effective and proportionate policy development, and can result in sub-optimal business decisions.

3 Technology as a key driver of innovation and change in financial markets[3]

The relentless development and deployment of new technologies will continue to have profound effects on markets at many levels. They will directly affect developments in HFT/AT and continue to fuel innovation in the development of new market services. And they will also help to drive changes in market structure.

New technologies are creating new capabilities that no human trader could ever offer, such as assimilating and integrating vast quantities of data and making multiple accurate trading decisions on split-second time-scales. Ever more sophisticated techniques for analysing news are also being developed and modern automated trading systems can increasingly learn from monitoring sequences of events in the market. HFT/AT is likely to become more deeply reliant on such technologies.

Future developments with important implications:

- *There will be increasing availability of substantially cheaper computing power, particularly through cloud computing:* those who embrace this technology will benefit from faster and more intelligent trading systems in particular.
- *Special purpose silicon chips will gain ground from conventional computers:* the increased speed will provide an important competitive edge through better and faster simulation and analysis, and within transaction systems.
- *Computer-designed and computer-optimised robot traders could become more prevalent:* in time, they could replace algorithms designed and refined by people, posing new challenges for understanding their effects on financial markets and for their regulation.
- *Opportunities will continue to open up for small and medium-sized firms offering 'middleware' technology components, driving further changes in market structure:* such components can be purchased and plugged together to form trading systems which were previously the preserve of much larger institutions.

3 For a more detailed review of the evidence reported in this section, see Chapter 2 in the Project's Final Report.

Three key challenges arising from future technological developments:

- *The extent to which different markets embrace new technology will critically affect their competitiveness and therefore their position globally:* The new technologies mean that major trading systems can exist almost anywhere. Emerging economies may come to challenge the long-established historical dominance of major European and US cities as global hubs for financial markets if the former capitalise faster on the technologies and the opportunities presented.
- *The new technologies will continue to have profound implications for the workforce required to service markets, both in terms of numbers employed in specific jobs, and the skills required:* Machines can increasingly undertake a range of jobs for less cost, with fewer errors and at much greater speed. As a result, for example, the number of traders engaged in on-the-spot execution of orders has fallen sharply in recent years, and is likely to continue to fall further in the future. However, the mix of human and robot traders is likely to continue for some time, although this will be affected by other important factors, such as future regulation.
- *Markets are already 'socio-technical' systems, combining human and robot participants. Understanding and managing these systems to prevent undesirable behaviour in both humans and robots will be key to ensuring effective regulation:* While this Report demonstrates that there has been some progress in developing a better understanding of markets as socio-technical systems, greater effort is needed in the longer term. This would involve an integrated approach combining social sciences, economics, finance and computer science. As such, it has significant implications for future research priorities.

4 The impact of computer-based trading on market quality: liquidity, price efficiency/discovery and transaction costs[4]

While the effect of CBT on market quality is controversial, the evidence available to this Project suggests that CBT has several beneficial effects on markets, notably:

- *liquidity, as measured by bid-ask spreads and other metrics, has improved;*
- *transaction costs have fallen for both retail and institutional traders, mostly due to changes in trading market structure, which are related closely to the development of HFT in particular;*
- *market prices have become more efficient, consistent with the hypothesis that CBT links markets and thereby facilitates price discovery.*

While the above improvements in market quality should not be overstated, they are important, particularly since they counter the belief that HFT provides no useful function in financial markets. Nevertheless, there are concerns relating to market quality which are worthy of mention.

A particular concern:

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. However, unlike 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. These factors, together with the ultra-fast speed of trading, create the potential for periodic illiquidity. The US Flash Crash and other more recent smaller events illustrate this increased potential for illiquidity.

A key message: regulatory changes in practices and policies will be needed to catch up to the new realities of trading in asset markets. However, caution needs to be exercised to avoid undoing the advantages that HFT has brought.

4 For a more detailed review of the evidence reported in this section, see Chapter 3 in the Project's Final Report.

5 Financial stability and computer-based trading[5]

The evidence available to this Project provides no direct evidence that computer-based HFT has increased volatility in financial markets. However, in specific circumstances CBT can lead to significant instability. In particular, self-reinforcing feedback loops, as well as a variety of informational features inherent in computer-based markets, can amplify internal risks and lead to undesired interactions and outcomes. This can happen even in the presence of well-intentioned management and control processes. Regulatory measures for addressing potential instability are considered in Section 7 of this Executive Summary.

Three main mechanisms that may lead to instabilities and which involve CBT are:

- nonlinear sensitivities to change, where small changes can have very large effects, not least through feedback loops;
- incomplete information in CBT environments where some agents in the market have more, or more accurate, knowledge than others and where few events are common knowledge;
- internal 'endogenous' risks based on feedback loops within the system.

The feedback loops can be worsened by incomplete information and a lack of common knowledge.

A further cause of instability is social: a process known as 'normalisation of deviance', where unexpected and risky events (such as extremely rapid crashes) come to be seen as increasingly normal, until a disastrous failure occurs.

6 Market abuse[6] and computer-based trading[7]

Economic research thus far, including the empirical studies commissioned by this Project, provides no direct evidence that HFT has increased market abuse[8]. However, the evidence in the area remains tentative: academic studies can only approximate market abuse as data of the quality and detail required to identify abuse are simply not available to researchers.

This Project has commissioned three empirical studies that find no direct evidence of a link between HFT and market abuse. The main focus of these studies is not on the measurement of market abuse during the continuous phase of trading, however. The Project has reviewed qualitative evidence on perceived levels of manipulation from various sources including interviews with traders and investors, the financial press, UK and international regulatory reports, submissions to regulatory consultations and large-scale surveys of market participants. A new survey of end users was also carried out by the Project[9].

This qualitative evidence consistently indicates high levels of concern. Claims of market manipulation using HFT techniques are reported by institutional investors such as pension funds and mutual funds in different countries. These claims are, in turn, widely relayed by the financial press. Even if not backed by statistical evidence, these perceptions need to be taken seriously by policy makers because, given that the true extent of abuse is not precisely known, it is perception that is likely to determine the behaviour of liquidity suppliers. High perceived levels of abuse may harm market liquidity and efficiency for all classes of traders.

The qualitative evidence mentioned above is not easy to interpret unambiguously. It is consistent with three different 'scenarios' that are not mutually exclusive:

- High frequency traders exploit their speed advantage to disadvantage other participants in financial terms.
- The growth of HFT has changed order flows in ways that facilitate market abuse by both slow and fast agents (for example, by making 'predatory trading' easier).

5 For a more detailed review of the evidence reported in this section, see Chapter 4 in the Project's Final Report.

6 Here the concern is with market abuse relating to manipulative behaviour, by which a market is temporarily distorted to one party's advantage. Abuse relating to insider trading is not considered here.

7 For a more detailed review of the evidence reported in this section, see Chapter 5 in the Project's Final Report.

8 A list of the studies commissioned by the Project may be found in Annex D of the Project's Final Report.

9 SR1 (Annex D refers).

- Other market developments concomitant with the growth in HFT, but not necessarily brought about by HFT growth, may have contributed to an increase in the perception or actual prevalence of abuse. Fragmentation of liquidity across trading venues is an example.

Regulators and policy makers can influence perceptions, even if definitive evidence on the extent of abuse will not be available to settle the debate.

- *Regulators can address the lack of confidence that market participants have in their ability to detect and prosecute abuse in HFT and fragmented markets.* While this may require significant investment in regulatory activity, if progress is made, both the perception and reality of abuse will be reduced; for abusers, even a perceived threat of being caught may be a powerful disincentive.
- *More statistical evidence on the extent of HFT manipulation most often described by institutional investors can be produced*[10]*.* This will help to correct or confirm perceptions. It will also be important in guiding regulatory action, as the three scenarios outlined above may have very different policy implications.

Detecting evidence of market abuse from vast amounts of data from increasingly diverse trading platforms will present a growing challenge for regulators.

To identify abuse, each national regulator will need access to international market data. Otherwise the market abuser can hide by transacting simultaneously in several separately linked markets. In the USA, the Office of Financial Research (OFR) has been commissioned by the Dodd-Frank Act to fund a financial data centre to collect, standardise and analyse such data. *There may be case for a similar initiative to be introduced in Europe.*

7 Economic impact assessments of policy measures[11]

A number of policies related to CBT are being considered by policy makers with the goals of improving market efficiency and reducing the risks associated with financial instability. This Project has commissioned a variety of studies to evaluate these policies, with a particular focus on their economic costs and benefits.[12] *The key conclusions are set out below.*

Policy measures that could be effective:

- Circuit breakers: *There is general support for these, particularly for those designed to limit periodic illiquidity induced by temporary imbalances in limit order books.* They are especially relevant to markets operating at high speed. Different markets may find different circuit breaker policies optimal, but in times of overall market stress there is a need for coordination of circuit breakers across markets, and this could be a mandate for regulatory involvement. New types of circuit breakers triggered by *ex-ante* rather than *ex-post* trading may be particularly effective in dealing with periodic illiquidity. However, further investigation is needed to establish how coordination could best be achieved in the prevailing market structure.
- Tick size policy: *This can have a large influence on transactions costs, market depth and liquidity provision.* The current approach of allowing each European trading venue to choose its own minimum tick size has merits, but this can result in a race to the bottom between venues. A uniform policy applied across all European trading venues is unlikely to be optimal, but a coherent overall minimum tick size policy applying to subsets of trading venues may be desirable. This coordinated policy could be industry-based, such as the one agreed to recently by the Federation of European Securities Exchanges (FESE) members.

10 'Quote stuffing' or order book 'layering' are obvious examples.

11 For a more detailed review of the evidence reported in this section, see Chapter 6 in the Project's Final Report.

12 A list of the economic impact assessments commissioned by the Project can be found in Annex D of the Project's Final Report.

Policy measures that are likely to be problematic:

- **Notification of algorithms:** The implementation of this, even if feasible, would require excessive costs for both firms and regulators. It is also doubtful that it would substantially reduce the risk of market instability due to errant algorithmic behaviour.
- **Imposing market maker obligations and minimum resting times on orders:** The former issue runs into complications arising from the nature of high frequency market making across markets, which differs from traditional market making within markets. Requirements to post two-sided quotes may restrict, rather than improve, liquidity provision. Similarly, minimum resting times, while conceptually attractive, can impinge upon hedging strategies that operate by placing orders across markets and expose liquidity providers to increased 'pick-off risk' due to the inability to cancel stale orders.
- **Order-to-execution ratios:** This would be a blunt instrument to reduce excessive message traffic and cancellation rates. While it could potentially reduce undesirable manipulative strategies, it may also curtail beneficial strategies. There is not sufficient evidence at this point to ascertain these effects, and so caution is warranted. Explicit fees charged by exchanges on excessive messaging, as well as greater regulatory surveillance geared to detect manipulative trading practices may be more desirable approaches to deal with these problems.
- **Maker-taker pricing:** The issue is complex and is related to other issues like order routing, priority rules and best execution. Regulatory focus on these related areas seems a more promising way of constraining any negative effects of maker-taker pricing than direct involvement in what is generally viewed as an exchange's business decision.
- **The virtual central limit order book (CLOB):** The introduction of competition between trading venues brought about by Markets in Financial Instruments Directive (MiFID) has resulted in more choices for investors and, in many cases, improved market quality, but it has also led to greater complexity and risk. The virtual CLOB it has created is still evolving and improving, but its current structure falls short of a single integrated market. This raises a number of issues for both individual exchange and market behaviour.
- **Constraining internalisation or, more generally, dark trading:** Off-exchange trading can be mutually beneficial for all parties involved, especially where large orders are involved. However, the trend away from pre-trade transparency cannot be continued indefinitely without detrimental effects on the public limit order book and price discovery. Constraining these activities within a range that does not adversely affect price discovery but does allow for beneficial trading is important but difficult. Evidence gathered from European markets is too limited to give satisfactory guidance.
- **Call auctions:** These are an alternative trading mechanism that would eliminate most of the advantage for speed present in modern electronic markets. They are widely used already in equity markets at open and close and following a trading halt, although no major market uses them exclusively. To impose call auctions as the only trading mechanism seems unrealistic and draconian. There are serious coordination issues related to hedging strategies that would make this policy undesirable.

Two words of caution: Whilst the above conclusions are consistent with the currently available evidence, further empirical study is desirable for some of the policy measures in particular. It should also be recognised that some of the above individual policy options interact with each other in important ways. For example, the presence or absence of circuit breakers affects most other measures, as does minimum tick sizes. Decisions on individual policies should not therefore be taken in isolation, but should take account of such important interactions[13].

13 See the Project's Final Report (Chapter 6, Section 6.12) and also the supporting evidence papers which were commissioned (Annex D of the Project's Final Report refers).

8 Computers and complexity

Over coming decades, the increasing use of computers and information technology in financial systems is likely to make them more, rather than less complex. Such complexity will reinforce information asymmetries and cause principal/agent problems, which in turn will damage trust and make the financial systems sub-optimal. Constraining and reducing such complexity will be a key challenge for policy makers. Options include requirements for trading platforms to publish information using an accurate, high resolution, synchronised timestamp. Improved standardisation of connectivity to trading platforms could also be considered.

However, there is no 'magic bullet' to address this issue. Policy makers will need an integrated approach based on improved understanding of financial systems. This will need to be achieved through:

- **Improved post-trade transparency:** The challenge of ensuring adequate dissemination and storage of trading data to enable market abuse to be identified provides an important example of where improvements need to be considered.
- **Analysis:** Making sense of disclosed information and developing a better understanding of the financial system will be critical. This implies the need to harness the efforts of researchers[14].

A further proposal that is sometimes made is that (various categories of) agents should only be allowed to hold or issue instruments which have been approved by the authorities in advance. This contrasts with the more common position that innovation should be allowed to flourish, but with the authorities retaining the power to ban the uses of instruments where they consider evidence reveals undesirable effects. The former stance, however, not only restricts innovation, but also such official approval may well have unintended consequences. Furthermore, the effectiveness of such official approval is debatable. Officials have no more, and probably less, skill in foreseeing how financial instruments will subsequently fare than credit rating agencies or market agents. Indeed, many, possibly all, of the instruments now condemned in some quarters as having played a part in the recent global financial crisis would, at an earlier time, have probably been given official approval.

A corrective step that could, and should, be taken is to simplify (electronic) financial systems by the application of greater standardisation, particularly in the form of accurate, high resolution, synchronised timestamps. CBT, operating on many trading platforms, has led to a vast expansion of data, which are often not standardised, nor easily accessible to third parties (for example, regulators and academics) for analysis and research. The relevant authorities should consider following the US example and establish a European Financial Data Centre to collect, standardise and analyse such data.

9 Conclusions – key priorities for action[15]

While the effects CBT on financial markets have been the topic of some controversy in recent years, analysis of the available evidence has shown that CBT has led to benefits to the operation of markets, notably relating to liquidity, transaction costs and the efficiency of market prices[16]. Against the background of ever greater competition between markets, it is highly desirable that any new policies or market regulation preserve these benefits.

However, this Project has also highlighted legitimate concerns that merit the close attention of policy makers, particularly relating to the possibility of instabilities occurring in certain circumstances, and also periodic illiquidity[17]. In view of the critical importance of financial markets for global growth and prosperity, the following suggests priorities for action:

14 See Section 9 of this Executive Summary.

15 For a more detailed review of this section, see Chapter 8 in the Project's Final Report.

16 See Section 4 of this Executive Summary.

17 See Section 5 of this Executive Summary.

A. Limiting possible future market disturbances:

A.1 European authorities[18], working together, and with financial practitioners and academics, should assess (using evidence-based analysis) and introduce mechanisms for managing and modifying the potential adverse side-effects of CBT and HFT. Section 7 of this Executive Summary sets out analysis of ten individual policy options, and provides advice on which are supported most by the available evidence. It is also important that such regulatory measures are considered together, not individually, in view of important interactions which may exist between some of them.

A.2 Coordination of regulatory measures between markets is important and needs to take place at two levels:

- *Regulatory constraints involving CBT in particular need to be introduced in a coordinated manner across all markets where there are strong linkages.*
- *Regulatory measures for market control must also be undertaken in a systematic global fashion to achieve in full the objectives they are directed at.* A joint initiative from a European Office of Financial Research and the US Office of Financial Research (OFR), with the involvement of other international markets, could be one option for delivering such global coordination.

A.3 Legislators and regulators need to encourage good practice and behaviour in the finance and software engineering industries. This clearly involves the need to discourage behaviour in which increasingly risky situations are regarded as acceptable, particularly when failure does not appear as an immediate result[19]. These recognise that financial markets are essentially complex 'socio-technical' systems, in which both humans and computers interact: the behaviour of computers should not be considered in isolation.

A.4 Standards should play a larger role. Legislators and regulators should consider implementing accurate, high resolution, synchronised timestamps because this could act as a key enabling tool for analysis of financial markets. Clearly it could be useful to determine the extent to which common gateway technology standards could enable regulators and customers to connect to multiple markets more easily, making more effective market surveillance a possibility.

A.5 In the longer term, there is a strong case to learn lessons from other safety-critical industries, and to use these to inform the effective management of systemic risk in financial systems. For example, high-integrity engineering practices developed in the aerospace industry could be adopted to help create safer automated financial systems.

B. Making surveillance of financial markets easier:

B.1 The development of software for automated forensic analysis of adverse/extreme market events would provide valuable assistance for regulators engaged in surveillance of markets. This would help to address the increasing difficulty that people have in investigating events.

C. Improving understanding of the effects of CBT in both the shorter and longer term:

C.1 Unlocking the power of the research community has the potential to play a vital role in addressing the considerable challenge of developing better evidence-based regulation relating to CBT risks and benefits and also market abuse in such a complex and fast-moving field. It will also help to further address the present controversy surrounding CBT. Suggested priorities include:

- *Developing an 'operational process map':* this would detail the processes, systems and interchanges between market participants through the trade life cycle, and so help to identify areas of high systemic risk and broken or failing processes.
- *Making timely and detailed data across financial markets easily available to academics, but recognising the possible confidentiality of such data.*

18 While several of these potential actions for stakeholders are framed within the European context, they will also be relevant to stakeholders in other parts of the world.

19 A term for behaviour which accepts increasingly risky situations in the absence of adverse effects is called 'normalisation of deviance'. See Section 5 of this Executive Summary.

C.2 The above measures need to be undertaken on an integrated and coordinated international basis in order to realise the greatest added value and efficiency. One possible proposal would be to establish a European Financial Data Centre.

In conclusion:

It is hoped that the analysis and arguments contained in the Foresight Final Project Report, together with over 50 commissioned evidence papers which underpin it, will assist policy makers, regulators and market practitioners in their current consideration of CBT. In this context, special thanks are due to the 150 or so leading and independent experts from over 20 countries who have been involved in this undertaking.

This Executive Summary and the underlying Project Report provide an independent view based on the best available science and evidence. They do not constitute the views or policy of the UK or any other government.

1 Introduction

1.1 The aim of this Project

This Project has two principal aims. First, looking out to 2022, it seeks to determine how computer-based trading in financial markets could evolve and, by developing a robust understanding of its effects, to identify potential risks and opportunities that this could present, notably in terms of financial stability[1] but also in terms of other market outcomes such as volatility[2], liquidity[3], price efficiency and price discovery[4]. Secondly, drawing upon the best available scientific and other evidence, the Project aims to provide advice to policy makers, regulators and legislators on the options for addressing those risks and opportunities.

1.2 Why the Project was commissioned

Computer-based trading (CBT)[5] has grown substantially in recent years, due to fast-paced technological developments and their rapid uptake, particularly in equity markets. For example, possibly 30% of the UK's equity trading volume is now generated through high frequency trading (HFT), while in the USA this figure is possibly over 60%[6]. CBT is therefore already transforming the ways in which financial markets operate.

Inevitably, substantial changes to the functioning of financial markets, actual or perceived, attract considerable attention, not least because of their potential impact on market confidence, the operation of businesses and the health and growth of economies.

HFT attracts particular controversy. There has been a continuing debate about the extent to which HFT improves or degrades the functioning of financial markets, and also its influence on market volatility and the risk of instabilities. Indeed, such trading has been implicated by some as a contributory factor in the Flash Crash of 6 May 2010 in which one trillion dollars temporarily evaporated from US markets[7]. However, the wider debate on HFT and CBT has been hampered by the availability of relatively little evidence, and a lack of analysis.

The controversy concerning the effect of CBT on financial systems has the potential to grow for several reasons. Relentless technological developments in both hardware and in trading algorithms, together with other drivers of change[8], continue to influence the structure of markets, inevitably creating winners and losers at all levels. They are also leading to increases in complexity as well as new dynamics, making markets ever harder to understand and to regulate, particularly in view of the rapid pace of change. These developments are fuelling an urgent need to gain a better understanding of a range of issues, most notably concerning their effect on systemic risk of financial instability and its management. However, other important concerns relate to the effect of CBT on the efficiency of markets which has particularly divided opinion; the evolving potential for market abuse, especially its

1 Financial market stability refers to the lack of extreme movements in asset prices over short time periods. A glossary of terms and a list of acronyms used in this Report can be found in Annex C.

2 Volatility is defined here as variability of an asset's price over time, often measured in percentage terms.

3 Liquidity is defined here as the ability to buy or sell an asset without greatly affecting its price. The more liquid the market, the smaller the price impact of sales or purchases. For a more detailed description see Section 3.3.1.

4 Price efficiency and price discovery – pricing is efficient when an asset's price reflects the true underlying value of an asset; price discovery refers to the market process whereby new information is impounded into asset prices.

5 See Box 1.1 for a definition of CBT.

6 Accurate estimates of the volume of high frequency trading are difficult to obtain, and in any case are contingent on the precise definition used. For estimates, see Kaminska (2011), Kaminska (2009) and http://www.tradersmagazine.com/news/high-frequency-trading-benefits-105365-1.html?zkPrintable=true Accessed: 3 September 2012.

7 Chapter 4 reviews the available evidence for the influence of HFT in the Flash Crash of 6 May 2010.

8 See Box 1.3 and Figure 1.1.

detection and regulation; and the relationship of CBT with dark pools and changing market institutions more generally.

For these reasons, CBT is currently attracting the close attention of policy makers and regulators worldwide. For example:

- In Europe, HFT has been placed near the top of the regulatory agenda, with a wide range of measures on CBT being debated in the European Union (EU) Parliament and the EU Commission within the Markets in Financial Instruments Directive (MiFID) II process. Indeed, some parties have mooted measures that could have very far-reaching implications for markets, as they could substantially constrain the future of CBT within Europe.
- In the USA, a number of measures are being proposed under the Dodd-Frank Act.
- A number of countries across the world, from Latin America to Asia, have adopted measures related to CBT as exchanges turn electronic.

Against this background of rapid change and the urgent needs of regulators, the relative lack of evidence and the prevalence of controversy over consensus is a major concern. Regulation that is not informed by evidence and analysis risks making matters worse rather than better. It was therefore decided to commission this Foresight Project to inform a broad audience of policy makers, regulators and other stakeholders around the world.

Box 1.1: Definition of computer-based trading

Computer-based trading (CBT) refers to the trading system itself. Financial institutions use CBT systems in a range of trading strategies, of which high frequency trading (HFT)[9] and algorithmic trading (AT) are two types. However, the use of a CBT system by a financial institution does not necessarily mean that it is a user of one or other of these strategies.

A useful taxonomy of CBT systems identifies four characteristics that can be used to classify CBT systems[10]:

1) CBT systems may 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).

2) CBT systems may adopt liquidity-consuming (aggressive) or liquidity-supplying (passive) trading styles.

3) CBT systems may engage in either uninformed or informed trading.

4) A CBT algorithm generates the trading strategy or only implements a decision taken by another market participant[11].

A more detailed definition may be found in Annex C.

9 For a definition of HFT please see Chapter 3.

10 DR5 (Annex D refers).

11 Please refer to Annex C for a comprehensive glossary and list of acronyms.

1.3 A robust, international and independent approach

This Foresight Project has taken a broad approach, drawing upon a wide range of disciplines, including economics, computer science, sociology and physics. In so doing, it has engaged with over 150 leading independent academics from more than 20 countries[12].

The Project has drawn upon a wider body of academic literature, but has also commissioned over 50 papers and studies, particularly where it was considered important to address gaps in the evidence base, or to collate and assess existing studies. A full list of commissioned work is presented in Annex D: all of this information is freely available from www.bis.gov.uk/foresight. Such work includes reviews of drivers of change, economic impact assessments of regulatory measures and the results of workshops held in Singapore, New York and London to gather the views of leading industry practitioners, economists and technologists on future scenarios. A sociological study and a survey of end users were also undertaken to understand the impact of computer trading on institutional investors. The commissioned work has been peer reviewed internationally by independent reviewers (except for certain workshop reports and a survey).

Throughout, the Project has been guided by a group of leading academics and senior industry practitioners. It has also benefited from the advice of an international group of high-level stakeholders, which has provided advice at critical stages of the Project. A full list of the 500 or so experts, academics and stakeholders who have been involved in the Project is provided in Annex A.

Box 1.2: An independent view

While the Project has been managed by the UK Foresight programme under the direction of Sir John Beddington, the UK Chief Scientific Adviser, its findings are entirely independent of the UK Government. As such, the findings do not represent the views of the UK or any other government, or the views of any of the organisations that have been involved in the work.

1.4 Project scope

The Project looks ten years into the future to take a long-term and strategic view of how CBT in financial markets might evolve, and how it might act within the context of other drivers of change, to affect a range of market functions including: financial stability, liquidity, price efficiency and discovery, transaction costs, technology and market abuse. While the future is inherently uncertain, major forces driving change can be identified. A key driver of change is technology and this report explores how technology and other drivers of change will interact to influence the development and impact of CBT. The possible consequences of a range of possible regulatory measures are also assessed.

Nine broad classes of drivers affecting financial markets have been considered. These were identified in three international industry workshops (held in London, Singapore and New York), and a workshop of chief economists in London. These drivers are briefly explained in Box 1.3 and are presented diagrammatically in Figure 1.1. In particular, this figure shows that all of the drivers affect the key aspects of market function, which in turn feed back to affect particular drivers[13].

While the work has taken a global view of drivers of change and markets, it has paid particular attention to the evolving use of CBT in Europe. However, much of the analysis will nevertheless be of interest to policy makers and markets in other parts of the world.

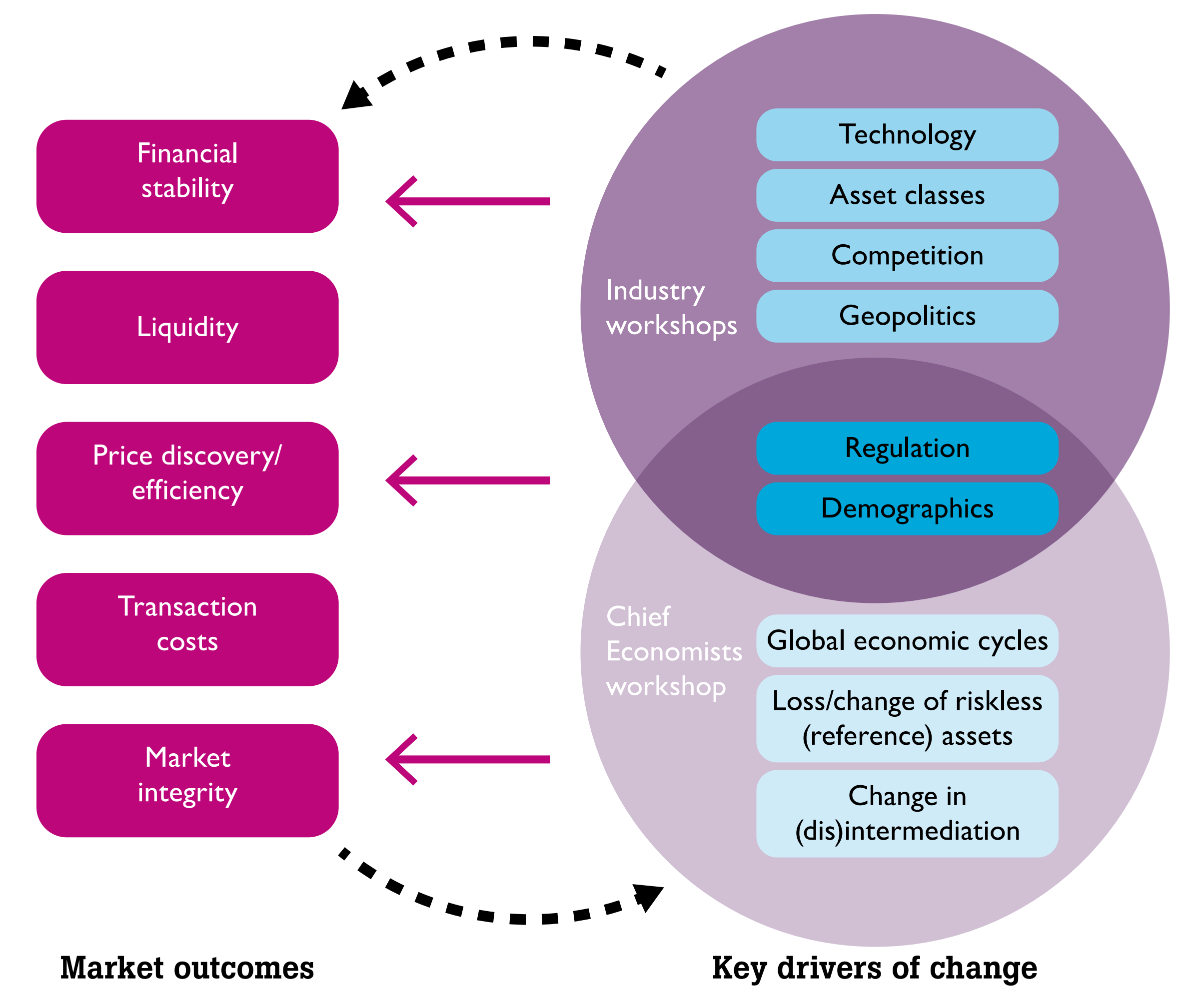
12 Please refer to Annex A for a list of all the individuals involved in this Project.

13 Please refer to Annex E for discussion on how drivers of change could play out in alternative future scenarios.

The analysis in this Report focuses particularly on high frequency and algorithmic trading. Its aim is to provide advice to policy makers who are taking decisions today on the regulation of CBT in markets – so that those decisions are well founded and are more likely to be robust to future uncertainties. However, in taking a ten year view, the Project also assesses what actions need to be implemented to address systemic issues such as financial instability, in the longer term. Some of these actions may, by their nature, take longer for policy makers to agree and implement.

The Project has taken an evidence-based approach wherever possible. However, in such a fast-moving field, gaps in both understanding and evidence are inevitable and these are highlighted in the text where appropriate. The most important gaps for informing regulation are identified in the concluding chapter. The Project has also explored the current perceptions of leading institutions about the impact of CBT in markets and importantly, it has evaluated the extent to which such perceptions are supported by the available evidence.

Figure 1.1: Key drivers of change



Box 1.3: Important future drivers of change

The following provides a brief summary of the nine broad classes of drivers identified during three international industry workshops (London, Singapore and New York), and a workshop of chief economists in London[14].

Regulation: Regulation will have an important and uncertain influence on financial markets, as both a driver and a consequence of financial market changes. As a driver, it may change the allocation of investment across assets and exchanges, and across institutions and investment models or strategies. Future regulation could be more or less coercive, informed by big data analytics, sophisticated models, heavy- or light-touch. There could be fragmentation at the global level, possibly linked to resurgent protectionism. Demand for regulation will tend to have an inverse relationship with levels of trust in the market.

Demographics: In the West over the next decade, the large number or people retiring will drive investment shifts, for example in the demand for retail rather than mutual funds, or for fixed income rather than equities.

Global economic cycles: The economic cycle appears to have been perturbed by the nature of the current recession. The dynamics of growth, employment, savings, trade, and leverage may return to previous cyclical behaviour or may instead follow a new pattern (prolonged recession, chaotic behaviour). Linked to this, global imbalances may persist or resolve. These factors will affect the demands placed on financial markets in terms of volume, asset classes and acceptable levels of risk and return. Global macroeconomic dynamics may also affect the process of globalisation and the relative importance of financial and 'real' markets.

Geopolitics: Growth rates over the next decade will powerfully influence the structure of future markets. A strong world economy will allow technological experimentation and new connections among geopolitical regions and groupings. A faltering economy or, worse, one in a tailspin, would be likely to lead to national retrenchment.

Technology[15]*:* This may lead to the creation of highly distributed trading platforms on which large numbers of individuals carry out transactions. Individual mobile phone handsets, possibly receiving live news and data feeds, may be used for trading; institutional trading strategies may also be influenced by communications on social networks. A new topology of either highly dispersed exchanges or of interlinked international exchanges could take shape.

Loss/change of riskless (reference) assets: The current global asset 'ecosystem' uses the return to riskless assets as a reference point for pricing risky assets. With sovereign debt now perceived to carry risk, this point of reference may be on the verge of disappearing. The behaviour of financial markets without a commonly recognised riskless asset is uncertain, and it is not clear whether a new common reference point will emerge. A connected issue is the link between sovereign debt and national currencies, and the role of the dollar as the global reserve currency.

Asset classes: Products focusing on levels of risk exposure rather than dividends may become more prominent; investment may shift from listed equity or derivatives towards synthetic products and spread-betting. These may focus on the state-dependent pattern or returns rather than ownership, and are likely to include more 'exotic' instruments. CBT may lead to the creation of new financial instruments for asset classes that are not currently directly traded using HFT or algorithmic tools.

14 See also Figure 1.1. More detailed discussion of these drivers can be found in the workshop reports (Annex D refers).

15 In view of the importance of technology as a key driver for computer-based trading, Chapter 2 provides a more detailed discussion. See also DR3 (Annex D refers).

Competition: Over and above technological changes, innovation in business models will shape competitive dynamics. Features analogous to Amazon's 'other products you may like' button may be introduced into institutional trading products. Market shares and returns will be driven by content, financial products and costs. Firms may unbundle services to generate more commissions or rebundle them to enhance investor lock-in. Exchanges are already increasing profits by proposing value-added components; they could increasingly focus on content-driven models.

Change in (dis)intermediation: Technological and financial market changes are altering both the size and role of intermediaries. The pace, direction and implications of these shifts will depend on whether such entities can operate across borders, the depth of funding that they influence and their impact on specific assets or investors. These developments are linked to CBT and HFT via the arbitrage role of intermediaries. They may be implemented via CBT/HFT by the continuum of CBT from algorithmic trading to HFT and by the degree to which the implications of CBT are different for the price trading and asset management functions of intermediaries.

1.5 Structure of the Report

This Report is comprised of eight chapters. Chapter 1 provides the rationale for undertaking the Project, sets out the aims and objectives, and its approach in terms of scope and content. In Chapter 2 technological developments are reviewed in detail, since this is a particularly important driver of change affecting CBT. Their recent impact is reviewed and technology advances likely in the next ten years for example, cloud computing, and custom silicon are explored. In Chapter 3 evidence on the impact of CBT on key indicators of market quality including liquidity, price discovery/efficiency and transaction costs are assessed. The chapter begins by examining the evidence for past effects, and then considers how impacts could change in the future, recognising that this will be contingent upon the mix of future regulatory measures in place.

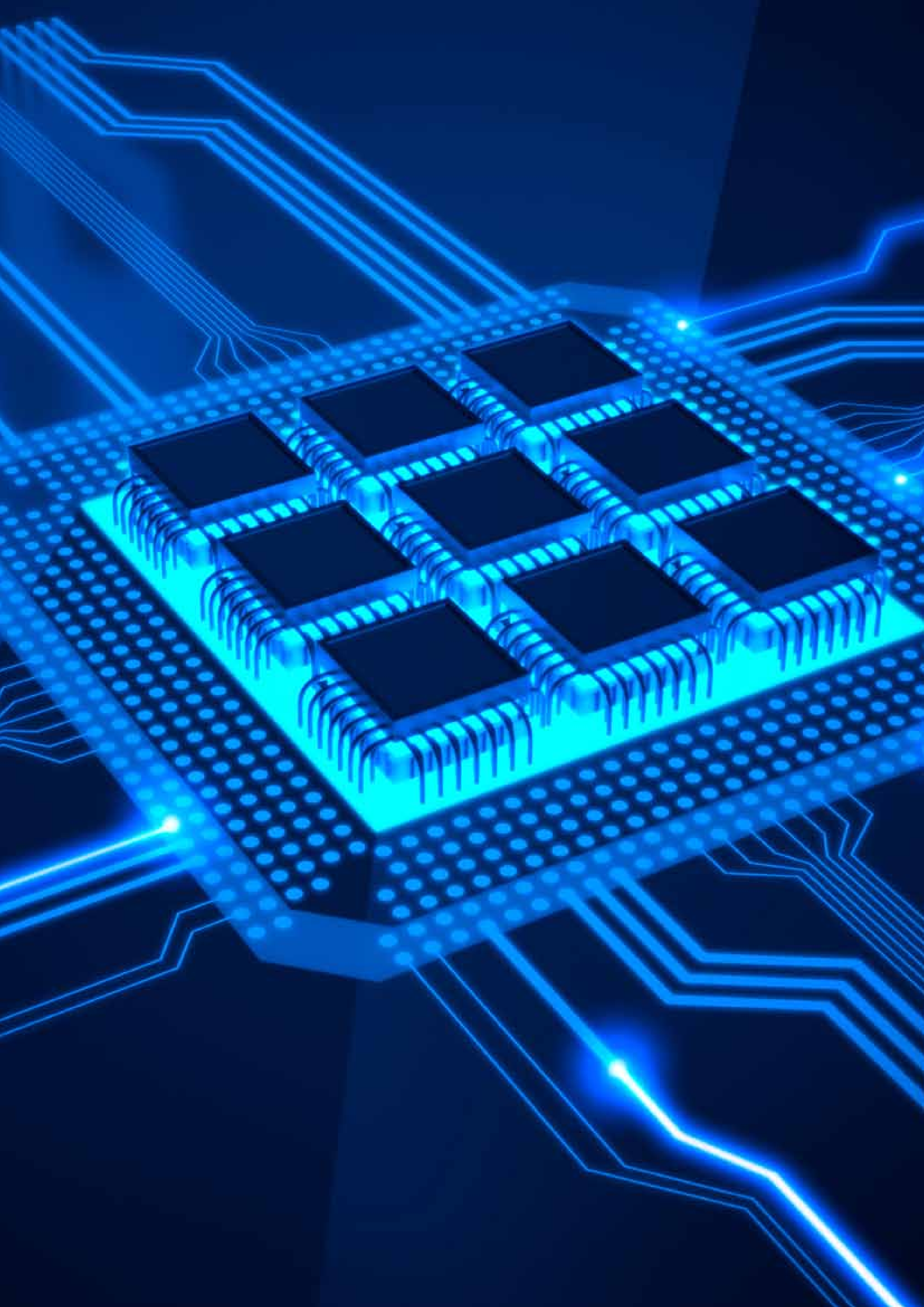
Chapter 4 examines the evidence for the impact of CBT on financial stability, evidence for past effects is reviewed and particular attention is given to the impact of self-reinforcing feedback loops in CBT which can amplify internal risks and lead to undesirable interactions and outcomes in financial markets. The concept of 'normalisation of deviance', where unexpected and risky events come to be seen as ever more normal, and its implications for financial stability is also explored.

The issue of market abuse is examined in Chapter 5 from economic, regulatory and user perspectives. It assesses the current impact of market abuse, and evidence on the perceptions of abuse using survey data commissioned by the Project. Consideration is also given to how the relationship between market abuse and CBT might evolve in the next ten years and possible courses of action to address the issue.

In Chapter 6 the potential economic impact of individual regulatory measures on stability, volatility and also liquidity, price discovery/efficiency and transaction costs are reviewed using a variety of new studies commissioned by the Project. Benefits as well as costs and risks are assessed. The measures are diverse ranging *inter alia* from notification of algorithms, circuit breakers, minimum tick size requirements and market maker obligations, to minimum resting times and periodic auctions.

The issue of how long term strategic factors, notably how CBT and HFT can affect trust and confidence in markets is discussed in Chapter 7. Particular emphasis is given to the role of rising complexity, enabled by information technology, in financial arrangements, transactions and processes in recent decades, and also the supply of credit, in influencing trust. It asks how can complexity be constrained and reduced and highlights the potential for a greater role for standards in addressing these issues.

Finally, Chapter 8 concludes the Report by drawing out the top level advice for policy makers, both for the short term and long term. In the latter case, priorities for research and better data collection are suggested.



2 The impact of technology developments

Key findings

Continuing advances in the sophistication of 'robot' automated trading technology, and reductions in cost 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.

2 The impact of technology developments

2.1 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. This trend is likely to have a corresponding 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 poorly understood. Moreover, the markets sometimes behave in unpredictable and undesirable ways. Few details are known of the connectivity network of interactions and dependencies in technology enabled financial markets. There is recognition that the current global financial network needs to be mapped in order to gain an understanding of the current situation. Such a mapping exercise would enable the development of new tools and techniques for managing the financial network and exploring how it can be modified to reduce or prevent undesirable behaviour[1]. New technology, new science and engineering tools and techniques will be required to help map, manage and modify the market systems of the future.

2.2 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 *et al.* (DR3)[2] relate the history of technology in the financial markets, briefly covering the 18th, 19th and 20th centuries, and then explore in more detail the rapid and significant changes which have occurred at the start of the 21st century[3].

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 outsmart their slower witted competitors. In the 19th century, communication of financially significant information by messengers on horseback 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 (CBT) systems meant that automated trading systems could start to perform functions previously carried out only by humans: computers could monitor the price of a financial instrument (for example, a share price) and 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 exacerbating 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), until by the late 1990s it was possible to buy, at no extra cost in real terms, computers over 50 times more powerful than those used in 1987. All this extra computer power could be put to use in implementing far more sophisticated processing for making investment decisions and for issuing structured patterns of orders to the markets.

By the end of the 20th century, as the real cost of computing continued to fall at a dramatic pace, the management of investment funds had become an increasingly technical field, heavily dependent

1 The need to map, manage, and modify the financial network is the central message of a speech *Rethinking the Financial Network* by Haldane (2009).

2 DR3 (Annex D refers).

3 For other accounts of the recent history of technology developments in the financial markets, the following three texts are particularly recommended: Angel *et al.* (2010), Gomber *et al.* (2011) and Leinweber (2009). For a recent discussion of high frequency trading, including interviews with leading practitioners, see Perez (2011) and Arnuk & Saluzzi (2012).

on computationally intensive mathematical models to manage portfolio risk (i.e. to 'hedge' the risk in the fund's holdings). Manual methods for hedging risk were used for decades before the deployment of electronic computers in fund management, but as the real cost of computing fell, and the speed and capacity of computers increased, so computers were increasingly called upon to calculate results that could guide the fund manager's decisions to buy or sell, and to go 'long' or 'short'. In this period, growing numbers of funds based their investment decisions on so-called statistical arbitrage (commonly abbreviated to 'stat arb'). One popular class of stat arb strategies identifies 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 will revert to its mean in due course. One of the simplest of these '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, it became possible to attempt increasingly 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 were considered more 'intelligent' 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 and 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' (AT) was born[4].

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 1980s, but typically involving 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 CBT 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 and fast automated execution via STP and DMA, means that, in recent years, it has become entirely commonplace for market participants to seek counterparties to a transaction electronically, identify a counterparty and then execute the transaction, all within a few seconds. For trades involving large quantities of a tradable instrument (so large that the volume of shares traded

4 One key aspect of modern automated trading systems is that they are designed to be highly autonomous: once switched on, they are intended by design to 'do their own thing', at high speed, with little or no human intervention at the level of individual trades. Although the computer running an algorithmic trading system can be switched out of action in an instant (for example, by severing its power cable), in many real systems it is much more desirable to order the algorithmic trader to go into its shutdown sequence, where it will do its best to sell off its portfolio of holdings as quickly as possible and with minimal losses before actually switching itself off. This means that while humans still bear ultimate responsibility for writing and managing algorithmic trading systems, humans are less and less involved in their minute-by-minute operation or control. An example of the owners of an automated trading system being unable to re-establish effective control occurred on 1st August 2012, when the US brokerage Knight Capital suffered massive losses as their automated trading system failed. See http://dealbook.nytimes.com/2012/08/03/trading-program-ran-amok-with-no-off-switch/?nl=business&emc=edit_dlbkam_20120806 Accessed: 4 September 2012.

is likely to significantly shift the price of the transaction – a phenomenon known as market impact), the computerised search for a counterparty is not often conducted in the open, on a national stock exchange, but instead takes place via the services of a private alternative electronic trading venue, a 'dark pool', which offers to match large volume counterparties in an anonymous or obfuscated fashion, to try to reduce market impact. Details of the transaction are only reported to the national exchange after it has been completed. Because many trading institutions see reducing market impact as a good thing, the popularity of dark pools has increased significantly in recent years.

The old 'vertically integrated' business model of investment banking is becoming increasingly fragmented. One effect of the European Union's (EU) Markets in Financial Instruments Directive (MiFID)[5] legislation was to create an ecosystem of small and medium-sized businesses offering 'middleware' technology components that could be purchased individually 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: with sufficient funding, 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)[6], where automated systems buy and sell on electronic exchange venues, sometimes holding a particular position for only a few seconds or less. 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 may only amount to a few pennies but if the HFT system is entirely automated, then it is a machine that can create a steady stream of profit every hour that the market is open. A recent study[7] indicated that the total amount of money extractable from the markets via HFT may be more modest than might be supposed: a few tens of billions of dollars in the US markets. Despite this, the low variation in positive returns from a well-tuned HFT system is an attractive feature and one that makes HFT an area of intense interest in the current markets. Of course, even if the total profits extractable via HFT really are limited, the downside risks, the total potential worst case losses and costs that might be incurred if HFT technology fails, could in principle be much larger. There is a concern that some CBT systems, like other novel trading systems in the past, could be making a steady stream of small profits but at the risk of causing very big losses if (or when) things go wrong. Even if each individual CBT system is considered to be stable, it is well known that groups of stable systems can, in principle, interact in highly unstable ways: the systemic risks of CBT are currently not well understood and are discussed in more detail in Chapter 4.

As the global financial markets became dependent on computers running automated trading systems and communicating with each other over optical fibre networks[8], the speeds of computation and of communication became two of the primary means by which competitive advantage could be gained and maintained. The effect of this in present day markets is discussed in the next section.

2.3 What are the key current technology developments?

2.3.1 From past to present

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 and the telecommunications networks that allow computers to communicate with each other. For the past

5 MiFID is an EU law that came into force in April 2004. In 2011, EU legislators began circulating drafts and revisions of a second version of MiFID, known as MiFID II.

6 Google Trends indicates that Google's users, worldwide, have only used the phrase 'high frequency trading' as a search term in the past four years, and 'algorithmic trading' only in the past six years. See www.google.com/trends Accessed: 17 September 2012.

7 Kearns *et al.* (2010).

8 The transmission speed of fibre-optic communications (i.e. via pulses of light beamed along glass fibres) is not quite equal to c, the speed of light in a vacuum. Transmission by line-of-sight microwave links through the atmosphere is slightly faster than fibre-optic communications and establishing microwave communications networks for HFT is currently an area of intense business activity. See Troianoivski (2012).

two decades, nearly all such firms used their own in-house information technology 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 (PC), of the sort available from high-street retailers, and the server computers would actually be constructed from several very high specification PCs, all located together and connected to each other in a single room; that room being the firm's 'server room' or 'data centre'.

As Cliff *et al.* (DR3) describe in some detail, the global information technology industry is currently undergoing a major shift towards cloud computing, where ultra large scale data centres (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 (See Figure 2.1). This greatly reduces the cost of high performance computing (HPC), and hence lowers barriers to entry for individuals or firms looking to use supercomputer scale HPC for the automated design and optimisation of trading systems. Rather than spending millions of dollars of capital expenditure on an in-house HPC data centre facility, it is now possible to obtain the same results by renting HPC from cloud computing providers for a few thousand dollars. 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.

Figure 2.1: Inside a cloud computing data centre, large numbers of individual computers are stacked in racks and any number of them can be remotely accessed via the internet as and when users require them. Typically, users pay only a small rental cost for each hour's access to each computer, so super-computer levels of processing power can be accessed at comparatively low cost.

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, offering huge increases in speed. This technology and the companies which have adopted it are discussed in more depth in DR3, which concludes 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. Technologists have therefore 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, status updates and 'tweets' on social media websites) and audio data (such as telephone calls, radio shows, podcasts and video sequences). This is an issue explored in depth by Mitra *et al.* (DR8), whose work is discussed in Section 2.3.2[9].

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 to expect that, for the next decade or so, the number of human participants in the market will remain significant. For most major markets in the USA, UK and mainland Europe, the proportion of computer-generated trades is estimated to be variously 30%, 60% or more[10]. What is clear is that the current markets involve large numbers of human traders interacting with large numbers of automated trading systems. This represents a major shift in the make-up of the markets and may well have affected their dynamics (see Chapters 1 and 7). Research that explores the interactions of humans and algorithmic trading systems is discussed in Section 2.3.3.

2.3.2 Automated analysis of market news and sentiment

Mitra *et al.* (DR8) consider 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 AT is. The traders in a market can be classified as one of two types: those who aim to profit by providing liquidity (so-called inventory traders) and those who aim to profit by trading on the basis of information (so-called informed or value motivated traders). 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 in any airport currency 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 informed/value motivated traders, make use of information in news stories and related discussion and analysis to form 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 differs 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 these technologies are rapidly increasing in sophistication.

In DR8, 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 (FX) 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 departments or scheduled earnings announcements by firms; unanticipated events are those such as news concerning major accidents, terrorist actions or natural disasters. Because of the effect that news events can have on the prices of financial instruments, major global

9 DR8 (Annex D refers).

10 Accurate estimates of the volume of high frequency trading are difficult to obtain, and in any case are contingent on the precise definition used. For estimates, see Kaminska (2011), Kaminska (2009) and http://www.tradersmagazine.com/news/high-frequency-trading-benefits-105365-1.html?zkPrintable=true Accessed: 3 September 2012.

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 reformatting or conversion from audio/video to text-based transcripts).

Because of these developments, 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 news analytics technology cannot 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.

2.3.3 Interactions between human and algorithmic trading systems

The global financial markets are now populated by two types of economic agents: human traders and 'software agents'. The latter are either algorithmic systems performing trading jobs that 10 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 have long been studied in the field known as experimental economics, and more recently the interactions between software agent traders in electronic markets have 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 surprisingly little scientific literature that explores heterogeneous markets, populated by both humans and algorithmic systems. In DR13[11], De Luca *et al*. provide a survey of the peer reviewed published research.

The first study to report any results was published in 2001[12] 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 Professor Jens Grossklags of the University of California at Berkeley, who used similar methods to explore different questions in studies published in 2003 and 2006[13]. Until 2011, the IBM experiments and Grossklags' work were the only three peer reviewed papers in the scientific literature that studied this topic. This is a startling gap in the research literature; one that has only very recently started to be filled.

In DR13, there is a detailed description and analysis of results from several preliminary experiments, conducted specifically for the Foresight Project. In DR25[14], this experimental work is replicated and extended. In the latter, 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 these preliminary experiments in DR13 and DR25 is that the previously reported outperformance of algorithmic trading systems in comparison to human traders may be due simply to the fact that computers can act and react faster than humans. This provides 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.

11 DR13 (Annex D refers).

12 Das *et al.* (2001).

13 Grossklags & Schmidt (2006).

14 DR25 (Annex D refers).

Johnson *et al.* (DR27)[15] argue that analysis of millisecond-by-millisecond US stock price movements between 2006 and 2011 suggests the existence of a step change or phase transition in the dynamics and behaviour of financial markets, in which human traders and automated algorithmic 'robot' trading systems interact freely. Above a particular time threshold, humans and algorithmic systems trade with one another; below the threshold, there is a sudden change to a market in which humans cannot participate and where all transactions are robot-to-robot. In addition to this, results in DR25 from a series of human-versus-robot experimental financial markets broadly support the hypothesis in DR27 concerning this change in market dynamics.

2.3.4 From the present to the future

Some key issues in today's markets look likely to remain vitally important in the future. For instance, cyber security will remain a core concern. Electronic attacks on the computer systems and communications networks of the global financial markets are always likely to be attractive to criminals. Furthermore, the widespread move towards greater 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. On the assumption that Einstein was right, this constraint is immutable.

The evolution of classes of algorithms is discussed in detail by Gyurkó in DR18[16]. Gyurkó concludes that the hunt for trading profits induces extreme competition, with any one firm's competitive advantage from a technology innovation being quickly eroded over time as competitors innovate and catch up, and hence market participants are forced to invest constantly in research and technology development. Recent academic research indicates that the next 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 these 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. Admittedly, knowledge about the functioning of financial markets populated entirely by humans is imperfect, even after hundreds of years; but present day markets with high levels of automation may be sufficiently different from human-only markets that new lessons need to be learned.

2.4 Technology advances likely in the next ten 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[17]. Nevertheless, several of the present day technology trends seem likely to remain significant factors over the next decade.

2.4.1 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 in 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 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

15 DR27 (Annex D refers).

16 DR18 (Annex D refers).

17 The arms race nature of adaptive trading technology development, and the steadily reducing levels of human involvement in major markets, means that algorithmic game theory may be a valuable tool in understanding systemic issues in automated financial markets. The well-known *Prisoners' Dilemma* game theory problem demonstrates that the stable outcome of rational interactions among competing self-interested entities is not necessarily the one that maximises the entities' rewards, i.e. the outcome may not be welfare enhancing.

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[18], 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 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 private data centres, but they do this for business critical operations and only a fraction of the capacity can be assigned to HPC uses. The ability to either extend existing in-house computing power by adding on externally provisioned cloud based resources, or to simply outsource all of the HPC provisioning to a cloud provider, opens up new possibilities that are only just being explored and lowers barriers to entry.

2.4.2 Custom silicon

General purpose, commercially available PCs have, over recent years, moved from being based on a single central processing unit (CPU) chip 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 towards so-called many-core computing, exploring the increased speed offered by using chips with many tens or hundreds of independent cores operating in parallel. This often involves using specialised processing chips originally designed for computer graphics processing. Furthermore, as noted above, the desire for ultra high speed processing of financial data has led to a move within the industry to replace general purpose computers 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 fit that user's needs).

Currently, the most popular technology of this type is a 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[19].

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) to only a few minutes from describing a trading algorithm in a high level programming language, to having that algorithm running on a parallel high speed computing array composed of a multitude of independent customised silicon chip processing elements. In DR3, it is suggested that this style of computer hardware is likely to be in wide use within ten years. 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.

18 Long-only macro trading is where a trader maintains a portfolio of holdings that are bought only in the expectation that their market value will increase (that is, the trader is taking *long* positions, as opposed to *short* positions where the trader would benefit if the market value *decreases*); and where the trader's alterations to the portfolio are driven by macroeconomic factors, such as national interest rates, which tend to alter relatively slowly, rather than the second-by-second fluctuations commonly exploited by high frequency traders.

19 A little more than a year after review DR3 was written, in May 2012, the trading technology provider company Fixnetix announced that they had delivered new products based on FPGA technology to a major bank in New York. The bank wishes to remain anonymous. Fixnetix stated that the FPGA-based hardware executed pre-trade compliance and 20 customisable risk checks "in nanoseconds". That is, the new hardware took less than one microsecond (less than one millionth of a second) to execute these pre-trade checks. See http://bit.ly/LN4seW Accessed: 17 September 2012.

2.4.3 Computer designed trading algorithms that adapt and learn

As De Luca *et al.* discuss at length in DR13, researchers have been studying and refining adaptive trading algorithms since the mid-1990s and, in 2001, IBM showed two such algorithms to be capable of outperforming human traders. Adaptive trading algorithms are automated trading systems that can learn from their experience of interacting with other traders in a market, improving their actions over time, and responding to changes in the market. Since the late 1990s, researchers have also studied the use of automated optimisation methods to design and improve adaptive trading algorithms. In automated optimisation, a vast space of possible designs is automatically explored by a computer program: different designs are evaluated and the best-performing design found by the computerised search process is the final output. Thus, new trading algorithms can be designed without the involvement of a human designer[20]. The use of these techniques in the finance industry looks likely to grow 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 may have had little or no human involvement in their design and refinement, and because they operate at truly superhuman speeds, 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 extremely difficult to predict or control. The concept of adaptive trading algorithms, whose behaviour may be difficult to understand, and whose dynamics in conjunction with other traders, whether automated or not, may be difficult to predict or control, will be quite disturbing. One natural instinct is to try to ban all such algorithms from one's own market. But that will not work, partly because the definition of what constitutes an adaptive algorithm is itself ill defined, and partly because markets will either exist, or can easily be set up, which will allow such trading practices. Once such markets exist, and they will, any market crashes or other disturbances in prices will reverberate directly back to the markets in which such algorithms are banned. Rather the need is to develop a culture and understanding which serves to limit such dangers.

Other studies, such as those discussed in DR25 and DR27, are needed to increase our understanding of market dynamics in the presence of automated trading systems that are adapting over time and whose designs are the result of computerised optimisation processes. Furthermore, as Cliff & Northrop describe at length in DR4[21], there are likely to be valuable lessons to be learned from fields other than finance or economics. In particular, sociological studies of how people interact with technology that is 'risky', in the sense that its safe operating limits are not fully understood in advance, have revealed that it is important to be mindful of the dangers of falling into a pattern of behaviour, a process called 'normalisation of deviance'. When a group of people engage in this process, they gradually come to see events or situations that they had originally thought to be dangerous or likely to cause accidents as less and less deviant or dangerous, and more and more normal or routine. This greatly increases the risk of major accidents. Normalisation of deviance is discussed in more depth in Chapter 4.

Normalisation of deviance can be avoided by adopting practices from what are known as high reliability organisations (HROs). Examples of HROs that have been studied include surgical teams, teams of firefighters and aircraft-carrier flight deck operations crews. In all such groups, safety is a critical concern, and it has been found that HROs have often independently developed common deviance monitoring processes involving careful post-mortem analyses of all procedures, even those in which no problems were apparent, conducted confidentially but with an internal atmosphere of openness and absence of blame. Adoption of HRO working practices and other approaches developed in safety-critical engineering (such as the use of predictive computer simulation models, as discussed in more depth in DR4, DR14, and DR17[22]) would help to mitigate possible ill effects of increased reliance on risky technology in the financial markets (for further discussion see Chapter 4).

20 Cliff (2009).

21 DR4 (Annex D refers).

22 DR14 and DR17 (Annex D refers).

2.5 Conclusions

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 new developments in computer technology.

Just as real physical robots revolutionised manufacturing engineering, most notably in automobile production, in the latter years of the 20th century, 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 decline in the number of front line traders employed by major financial institutions in the past few years as automated systems have been introduced is likely to continue over the next few years.

Nevertheless, there may be increased demand for developers of automated trading systems, and for designers of customised computer hardware that runs at high speed and with low latencies. It is most likely that the skills needed in these designers and developers will be those learned in advanced undergraduate and postgraduate degree courses. From a UK perspective, serious investment in coordinated programmes of research and development (for example, funded by the UK research councils or the UK Technology Strategy Board)[23] could help to secure the future 'talent base' (i.e. the pool of trained scientists and engineers that have skills appropriate for work on advanced trading algorithms and hardware).

The increased reliance on CBT is certainly not without its risks. Sudden market crashes (or sudden bubbles) in the prices of financial instruments can occur at greater speed, with chains of events proceeding at a much faster pace than humans are naturally suited to deal with. Furthermore, the globally interconnected network of market computer systems arguably means that an adverse event in one market now has greater potential to trigger a wave of contagion that could affect markets around the world (see Chapter 4 for further discussion of systemic risks of CBT). Equally, natural hazards in the form of floods or electrical storms can incapacitate data centre telecommunications networks; an interplanetary coronal mass ejection (a huge burst of plasma and electromagnetic energy from the sun, causing a geomagnetic storm when the ejection reaches the Earth) could seriously disrupt or disable the electrical power systems and radio communications of an entire city or larger area. Ensuring the resilience of critical computer and communications systems in the face of such major natural hazards, and in the face of attack from cyber criminals, must always remain a priority.

Even in the absence of such exogenous hazards, there are serious issues to be addressed in dealing with major endogenous risks, such as destabilising systemic internal feedback loops, and the pernicious effect of normalisation of deviance in risky technology systems, both of which are discussed further in Chapter 4. In addressing issues of systemic risk and financial stability, it is particularly important for regulators to develop and maintain the capacity to analyse extremely large data-sets. The challenge of dealing with 'big data' is already severe enough for any single major financial trading venue (for example, an exchange or dark pool): the number of active HFT systems and the rate at which they generate (and cancel) orders means that the venue's order book showing the best bid and offer for a particular stock may need to be updated thousands of times per second. A slight shift in the price of that stock can then cause the prices of many tens or hundreds of derivative contracts, such as options or exchange traded funds, to need to be recalculated. For regulators, the issue is made even more difficult as they are required to deal with aggregate market data generated simultaneously by multiple trading venues, as part of their regulatory role to oversee fair and orderly market systems, to identify policy violations and to monitor the need for policy revisions. Well designed mechanisms to manage the risk of volatility are one means of reducing the effects of risk in today's and future markets.

23 The UK does already have significant investment in relevant PhD-level research and training, with doctoral students from various initiatives such as the UK Doctoral Training Centre in Financial Computing and the UK Large-Scale Complex IT Systems Initiative (www.financialcomputing.org and www.lscits.org, respectively) working on issues of direct relevance to the topics addressed in this chapter. Despite this, there has been comparatively little investment in coordinated programmes of relevant academic research among teams of post-doctoral level research workers; that is, major research projects, requiring more effort than a single PhD student working for three years do not seem to have been prioritised.

On the basis of the evidence reviewed in this Chapter, it is clear that both the pace of development of technology innovations in financial markets and the speed of their adoption look set to continue or increase in the future. One important implication of these developments is that trading systems can today exist anywhere. Emerging economies, such as those of Brazil, Russia, India and China, may capitalise on the opportunities offered by the new technologies and in doing so may possibly, within only a few decades, come to rival the historical dominance of major European and US cities as global hubs for financial trading.

The next chapter builds on the information provided here to assess the impact of CBT on liquidity, price efficiency/discovery and transaction costs.


-0.73 0.05%
8.49 0.48%
4.83 0.27%
-0.54 0.03%
0.01 0.00%
1.34 0.17%
7.63 0.23%
-4.89 0.16%
-0.73 0.05%
8.49 0.38%
4.83 0.27%
-0.54 0.13%
0.01 0.10%
1.34 0.17%
5.63 0.23%
-6.89 0.06%
-0.73 0.15%
8.49 0.48%
778.33 1.34 0.17%
4443.87 7.63 0.23%
2916.60 -4.89 0.16%
1112.11 -0.73 0.05%
1787.63 8.49 0.38%
1791.97 4.83 0.27%
1295.09 -0.54 0.13%
767.89 0.01 0.10%
700.33 1.34 0.17%
443.83 5.63 0.23%
416.60 -6.89 0.06%
412.11 -0.73 0.15%
795.63 8.49 0.48%
791.97 4.83 0.27%
795.09 -0.54 0.53%
767.89 0.01 0.00%
778.33 1.34 0.17%
2443.83 5.63 0.23%
2416.60 -6.89 0.06%

3 The impact of computer-based trading on liquidity, price efficiency/discovery and transaction costs

Key findings

The past ten years have been a difficult period for investors in European and US financial markets due to low growth and a succession of crises. Nevertheless, computer-based trading has improved liquidity, contributed to falling transaction costs, including specifically those of institutional investors, and has not harmed market efficiency 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, leading to increased participation in high frequency trading, limiting its profitability.

However, there are some 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 need to be addressed.

3 The impact of computer-based trading on liquidity, price efficiency/discovery and transaction costs

3.1 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, programmed to meet particular trading requirements, organise orders to trade both temporally across the trading day and spatially across markets. High frequency traders use ultra-fast computing systems 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 to predict their costs from using alternative strategies.

What is particularly striking is that virtually all of these innovations have occurred within the past ten years[1]. In this short interval, the market ecology has changed, with markets evolving from traditional exchange-based monopolies to networks of computer-linked venues[2]. 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 chapter is to assess the evidence on 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. The implications of these recent and potential developments for policy makers and regulators are discussed in Chapter 8.

3.2 Determining the impact of computer-based trading on liquidity, price efficiency/discovery and transaction costs

Determining the impact of technology on market quality (the general term used to describe the liquidity, transaction costs 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 (HFT) is a case in point. HFT was virtually unknown five years ago, yet it is estimated that high frequency traders in the USA at times participate in 60% or more of trades in equities and futures markets[3].

The US Securities and Exchange Commission (SEC) 2010 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, machine learning and other quantitative techniques. Yet, even within this general description, the SEC notes the difficulty in characterising what HFT 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

1 See DR5 for discussion of the origins and growth of computer-based trading (Annex D refers).

2 See DR6 for discussion (Annex D refers).

3 Accurate estimates of the volume of high frequency trading are difficult to obtain, and in any case are contingent on the precise definition used. For estimates, see Kaminska (2011), Kaminska (2009) and http://www.tradersmagazine.com/news/high-frequency-trading-benefits-105365-1.html?zkPrintable=true Accessed: 3 September 2012.

> 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)[4].

Despite the lack of clarity as to the exact meaning of HFT, there is little disagreement regarding its importance in markets. Many high frequency traders act as market makers by placing passive limit orders onto electronic order books[5]. These passive orders provide the counterparty 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 high frequency traders are involved in a large percentage of market volume. Estimates of HFT involvement in US equity trading can be as high as 60%, with estimates of HFT involvement in European equities markets ranging from 30–50%[6]. Estimates of HFT in futures foreign exchange (FX) markets are in a similar range. Profits of high frequency traders in the USA in 2009 have been estimated at $7.2 billion[7], although others argue that the actual figure is much lower[8]. By comparison, the total value of electronic order book share trading on the NYSE and NASDAQ during 2009 was in excess of $30 trillion[9].

There are some important differences between such high frequency market making and its traditional specialist-based counterpart. HFT market makers rely on high speed computer linkages (often achieved by co-locating their 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, microseconds) and some operate with very low levels of capital, whereas specialists in traditional markets have obligations to stand ready to buy and sell. HFT 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[10]. This behaviour raises the spectre of periodic illiquidity.

The debate surrounding HFT 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[11]. Paul Krugman represents the contrarian view of HFT:

> 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

4 See US Securities and Exchange Commission (2010).

5 See DR10 and DR12 (Annex D refers).

6 Accurate estimates of the volume of high frequency trading are difficult to obtain, and in any case are contingent on the precise definition used. For estimates, see Kaminska (2011), Kaminska (2009) and http://www.tradersmagazine.com/news/high-frequency-trading-benefits-105365-1.html?zkPrintable=true Accessed: 3 September 2012.

7 See http://online.wsj.com/article/SB10001424053111904253204576510371408072058.html Accessed: 6 September 2012.

8 Kearns *et al.* (2010).

9 This is the total number of shares traded through the exchange's electronic order book multiplied by their respective matching prices. See World Federation of Exchanges: http://www.world-exchanges.org/statistics/annual Accessed: 17 September 2012.

10 It should be noted that traditional specialists (before the advent of CBT) also typically avoided holding large inventory positions. The number of specialists obliged to make markets has dwindled in the face of new competition from CBT. 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 trade on 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.

11 High frequency traders generally use proprietary data feeds to get information on the state of the market as quickly as possible. In the USA, this means that they receive information before it is delivered over the consolidated tape, raising issues of fairness and potential harmful effects on the cost of capital. See US Securities and Exchange Commission (2010).

> 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[12].

Rhetoric aside, while the issues surrounding computer-based trading (CBT), 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[13].

3.3 Past: what has been the impact of computer-based trading on market quality in recent years?

In this section the evidence on the effects on market quality of CBT is surveyed with emphasis being on the highest quality peer reviewed work. However, CBT is a relatively recent phenomenon and there is not much published research available. The field is evolving and so some caution is required in drawing conclusions. By way of comparison, the issue of whether smoking has negative health outcomes has concluded with almost universal agreement based on extensive, careful research. As the research on CBT is so recent, it has not yet achieved that level of agreement. Some discussion of methodological issues is presented in Box 3.1.

Box 3.1: Methodology

Some of the methodological issues relating to measurement and statistical inference considered in this chapter are discussed briefly here.

Measurement of HFT

There are two approaches to the measurement of CBT or, more specifically, HFT:

1. To use proxies (like message traffic) that are related to the intensity of trading activity. An example of this methodology is given in review DR23[14]. One issue with this approach is that it can be a poor guide to the presence or absence of HFT. For example, in DR23, the trading venues with the highest message traffic were in China, where HFT penetration is relatively light compared with London or Frankfurt. The message traffic in the Chinese case is generated by a large number of lower frequency retail traders.
2. To use (where available) trader identities which allows for classification through their observed strategies. Examples using US data can be found in Brogaard (2010)[15] and Hendershott *et al.* (2011)[16]. In Europe, Menkveld (2012) studied data from Chi-X/Euronext and identified an HFT market maker by their trading strategy. The driver review DR21[17] uses the Financial Services Authority (FSA) database and also involves directly classifying traders based on their observed behaviour.

12 Krugman (2009).

13 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 USA saw dramatic regulatory change in the guise of MiFID and Reg NMS, respectively. Consequently, care must be taken before ascribing most of the change in the market's behaviour to particular technological innovations.

14 DR23 (Annex D refers).

15 Brogaard (2010).

16 Hendershott *et al.* (2011).

17 DR21 (Annex D refers).

Measurement of market quality

A second issue concerns how to measure market quality: namely liquidity, price efficiency and transaction costs. Some of the evidence reviewed in this Report has used quoted bid-ask spreads and depth at the best quote, which are based solely on the order book. These methods have been criticised by some as representing a mirage of liquidity[18]. They also do not take account of market impact, which is an important component to transaction costs. However, other evidence is also reviewed where measures are based on effective and realised spreads, which make use of the order book and the transaction record, taking into account the impact of a trade. This counters the mirage problem to some extent. These approaches are widely used by academics and practitioners. The SEC mandates that trading venues in the US report and disseminate these measures monthly through their Rule 605[19]. This rule states:

> The term average *effective spread* shall mean the share-weighted average of effective spreads for order executions calculated, for buy orders, as double the amount of difference between the execution price and the midpoint of the consolidated best bid and offer at the time of order receipt and, for sell orders, as double the amount of difference between the midpoint of the consolidated best bid and offer at the time of order receipt and the execution price.

This impounds price movements after the trade (including the price impact due to the information in the trade). This cost can be interpreted as the profit realised by the other side of the trade, assuming they could lay off the position at the new quote midpoint. The realised spread and effective spread can be combined to measure the price impact, specifically:

Price impact = (effective spread – realised spread) / 2

Other liquidity measures include daily traded volume and turnover. The so-called Amihud measure of illiquidity, which measures the amount of price change per unit of trading volume, is also widely used and is described in review DR1[20].

Variance ratios are standard ways to measure price efficiency and are discussed in the review DR12[21]. They compare, for example, the variance of weekly returns with (five times) the variance of daily returns. If markets were fully efficient this ratio should be one, and any departure from this is evidence of market inefficiency. Low variance ratios imply negative dependence, meaning that prices tend to overreact, whereas high variance ratios imply positive dependence, meaning that prices tend to follow short-term trends.

18 Haldane (2011).

19 See Hasbrouck (2010).

20 DR1 (Annex D refers).

21 DR12 (Annex D refers).

Statistical issues

A key issue in identifying the consequences of HFT for market quality is *endogeneity*. That is, property *x* may be both a cause of HFT activity and a consequence of HFT activity. For example, HFT may cause some volatility, but some volatility may cause HFT to increase since this offers many profitable trading opportunities. The econometric methods available to identify the direction of causation or rather to control for endogeneity are as follows.

In a panel data context, where the variable of interest and the amount of HFT per unit per day is measured, *differences in differences* may be taken; that is, the difference between treated and untreated before and after the treatment (where treatment means the active presence of HFT) are compared. This approach eliminates common time-specific and firm-specific effects that may contaminate our view of the effects. This approach allows for a limited type of endogeneity but is very simple to apply and widely used in applied microeconomics.

A second approach is to use an *instrumental variable*; that is, a variable that is related to the input variable (for example, HFT activity) but unrelated to the output variable (for example, market quality) except directly through the input variable HFT. For example, the review DR21 used latency upgrade events on the London Stock Exchange (LSE). Analysis in review DR23[22] used the time at which an exchange first adopted automation. The main problem with this approach is finding credible instruments, that is, those that are not related directly to the output variable.

3.3.1 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, raises a variety of questions. Does it matter how much a trader wants to trade? Does time, or how long it takes to execute a trade, also matter? Will 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[23]. 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 quoted bid-ask spreads (tightness), the number of orders resting on the order book (depth) and the price impact of trades (resilience). These order book measures may not provide a complete picture since trades may not take place at quoted prices, and so empirical work considers additional measures that take account of both the order book and the transaction record. These include the so-called effective spreads and realised spreads, which are now widely accepted and used measures of actual liquidity[24]. Most of the studies reported below use these metrics.

In current equity markets, a common complaint is that liquidity is transient, meaning that orders are placed and cancelled within a very short time period and so are not available to the average investor. For example, the survey commissioned by this Project[25] reports that institutional buy-side investors believe that it is becoming increasingly difficult to access liquidity and that this is partly due to its

22 DR23 (Annex D refers).

23 Kyle (1985) and O'Hara (1995) for discussion.

24 In the USA, information about effective spreads must be provided by market centres that trade national market system securities on a monthly basis as a result of Exchange Act Rule 11Ac1-5.

25 See *Survey of end-users* in SR1 (Annex D refers).

fragmentation on different trading venues, the growth of 'dark' liquidity[26] and the activities of high frequency traders. 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's 'Guerrilla', Goldman Sachs' 'Stealth', or ITG's 'Dark', are designed to find liquidity without revealing the trading intentions, or even the existence, of the order submitter. These complications are not without costs to investors: the commissioned survey SR1 reports buy-side concerns that in order to access liquidity they have had to 'tool up' on expensive hardware and software or become more reliant on broker solutions that embed these costs.

The main empirical question is whether CBT (either in the guise of algorithmic trading (AT) or high frequency activity more generally) is associated with a decrease or increase in liquidity during regular market conditions. An equally important question relates to whether such trading exacerbates liquidity problems in situations of market stress.

There are a number of studies that try to identify CBT and its consequences on the order book and transactions. Hendershott *et al.* (2011) use the automation of the NYSE quote dissemination as an implicit experiment to measure the causal effect of AT on liquidity[27]. In 2003, the NYSE began to phase in the auto-quote system, which empowered CBT, initially for six large active stocks and then slowly over the next five months to all stocks on NYSE. They found that this change narrowed spreads, which was interpreted as increased AT improving liquidity and reducing adverse selection. The evidence was strongest for large stocks. Another study by Chaboud *et al.* (2009) also reports results on liquidity in the Electronic Broking Services (EBS) exchange rate market[28]. They found that, even though some algorithmic traders appear to restrict their activity in the minute following macroeconomic data releases, they increased 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[29]. 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. These are likely to be parts of a dynamic strategy adopted by such traders. 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 reducing short-term volatility.

Brogaard (2010) also investigated the impact of HFT on market quality in US markets[30]. High frequency traders were found to participate in 77% of all trades and tended to engage in a price-reversal strategy. There was no evidence to suggest that high frequency traders were withdrawing from markets in bad times or engaging in abnormal front-running of large non-HFT trades. High frequency traders demanded liquidity for 50.4% of all trades and supplied liquidity for 51.4% of all trades. They also provided the best quotes approximately 50% of the time.

Turning to Europe, Menkveld (2012) studied in some detail the entry of a new high frequency trader into trading on Dutch stocks at Euronext and a new market Chi-X in 2007 and 2008[31]. 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. All the trader's earnings arose from passive orders (liquidity supply). He also found that the bid-ask spreads were reduced by about 30% within a year when compared with Belgian stocks that were not traded by the HFT entrant.

26 Regarding the FTSE All-Share index for the week beginning 6th August, 2012: 47.9% of trading volume occurred on lit venues, 46.1% occurred on OTC venues, 3.8% was reported on dark pools, and 2.3% on Systematic Internalizers. See FIDESSA http://fragmentation.fidessa.com/fragulator/ The categories: lit, OTC, dark, and SI are defined by MiFID and they contain some overlap, but the only category that is fully transparent is the lit class. Even the lit class of venues such as the LSE allows so-called iceberg orders that are only partially revealed to the public. Accessed: 17 September 2012.

27 Hendershott *et al.* (2011), *op. cit.*

28 Chaboud *et al.* (2009).

29 Hasbrouck & Saar (2011).

30 Brogaard (2010).

31 Menkveld (2012).

There are also studies reporting trends in liquidity without specifically linking them to algorithmic or HFT. Castura *et al.* (2010) investigated trends in bid-ask spreads on the Russell 1000 and 2000 stocks over the period 2006 to 2010[32]. 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 *et al.* (2010) show a slow decrease in the average spread for S&P500 stocks over the period 2003 to 2010 (subject to some short-term up-side fluctuations in 2007/2008)[33]. 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.

Two commissioned studies provide evidence for UK markets. Friedrich 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 have 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 fivefold 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 as 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.

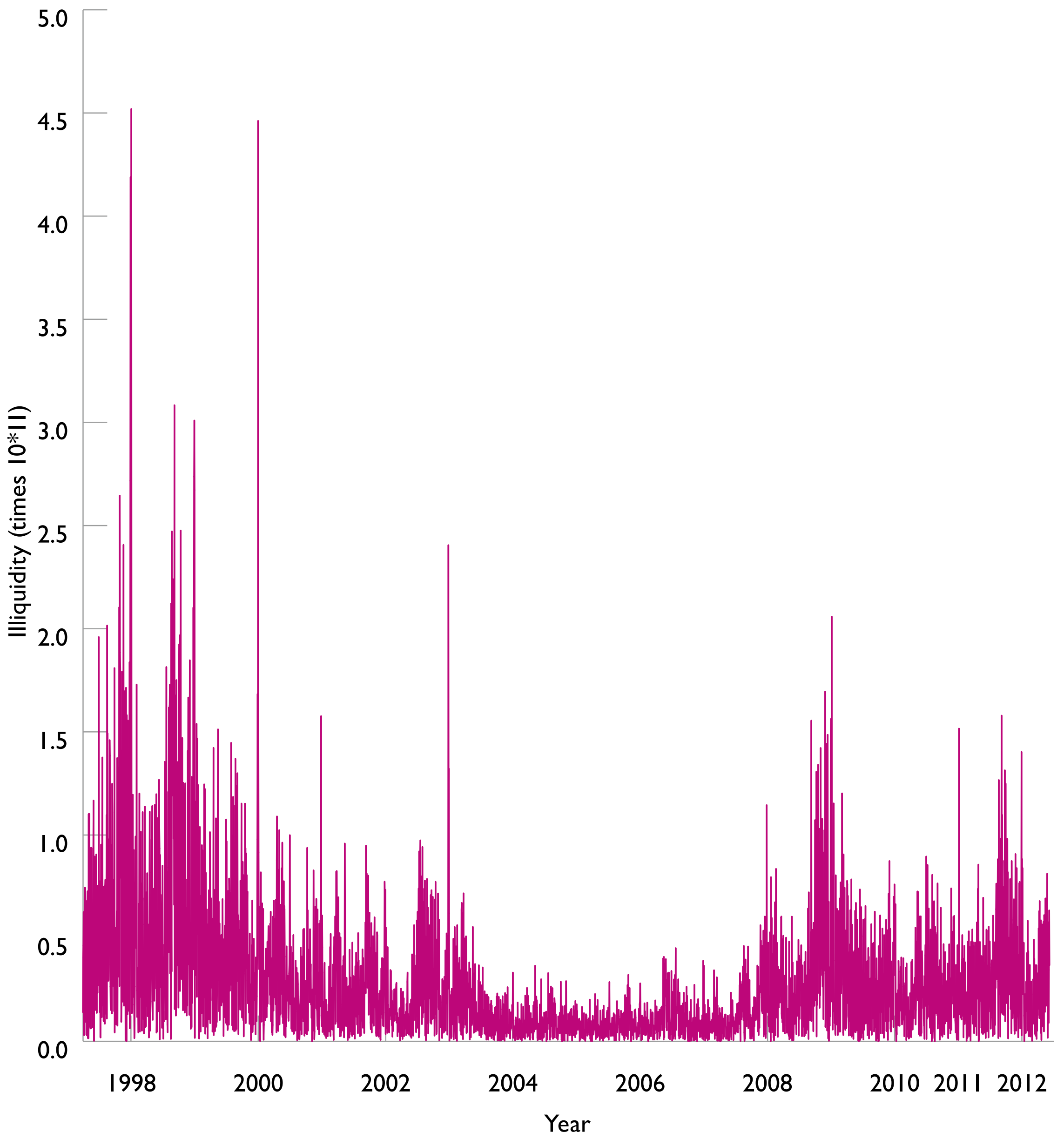
Linton (DR1) measures the daily illiquidity of the FTSE All-Share index using a low-frequency measure – the absolute return to unit of volume[34]. He finds this measure of illiquidity has varied considerably over the past ten years, first declining then rising during the financial crisis before falling again. Figure 3.1 shows an updated plot of the series up to June 2012.

32 Castura *et al.* (2010).

33 Angel *et al.* (2010).

34 The FTSE All-Share Index, (originally known as the FTSE Actuaries All Share Index,) is a capitalisation-weighted index, comprising around 1,000 of more than 2,000 companies traded on the London Stock Exchange. As at June 2011, the constituents of this index totalled 627 companies (source: FTSE All-Share fact sheet). It aims to represent at least 98% of the full capital value of all UK companies that qualify as eligible for inclusion. The index base date is 10 April 1962, with a base level of 100. The Amihud illiquidity measure is described in Box 3.1; it measures percentage price change per unit of trading volume.

Figure 3.1: FTSE All-Share illiquidity 1997-2012



Source: FTSE

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 study is that the macroeconomic crises of the past five years have been the main driving force for the liquidity and volatility of UK equity markets. If HFT has a role to play in the large swings in market conditions, it is relatively small and insignificant in comparison with the huge negative effects of the banking and sovereign debt crises that happened during this period. The level of the FTSE All-Share index has oscillated wildly throughout the past decade, as can be seen from Figure 3.2. Any medium term investor who was in the market during this period would have faced the prospect of having the value of his investments cut in half on two separate occasions and, if he held the market throughout the whole period, he would have ended up more or less where he started in nominal terms.

Figure 3.2: FTSE All-Share Index 1997-2012

Source: FTSE

The fragmented nature of liquidity was remarked on by a number of contributors to the survey of end users commissioned by the Project and reported in SR1[35]. In Europe, this fragmentation was facilitated by the Markets in Financial Instruments Directive (MiFID) regulation of November 2007, which permitted competition between trading venues. This market structure has been a seedbed for HFT, which has benefited from the competition between venues through the types of orders permitted, smaller tick sizes, latency and other system improvements, as well as lower fees and, in particular, the so-called maker-taker rebates. It is almost impossible to access liquidity effectively on multiple venues without sophisticated computer technology and, in particular, smart order routers (SOR).

35 SR1 (Annex D refers).

There are a number of studies that provide evidence on the effects of competition between trading venues. O'Hara and Ye (2011) find that stocks with more fragmented trading in the USA had lower spreads and faster execution times[36]. Gresse (DR19)[37] investigates whether this liquidity fragmentation has had a positive or negative effect on market quality in Europe. She examines this from two points of view. First, from the perspective of a sophisticated investor who has access to SORs and can access the consolidated order book. Second, from the point of view of an investor who can only access liquidity on the primary exchange. She compares FTSE100, CAC40 (French large caps) and SBF 120 (French mid caps) stocks across pre-and-post MiFID periods. She finds that fragmentation has generally had positive effects on spreads both locally and globally, especially for the larger cap stocks, whereas for the smaller cap SB120 the improvements were minor. For example, from October 2007 to September 2009, average (global) spreads on FTSE100 stocks fell from 9.21 to 5.43 basis points (bps), while the local spreads on average fell to 7.07, somewhat less. Spreads narrowed far less dramatically for the mid caps of the SBF 120.

These summary statistics are supplemented by a panel data analysis that controls for other variables that might affect outcomes and, in particular, acknowledges the background of the financial crisis and its potential effect on market quality. The regression study generally confirms the findings, although the improvements of the SBF 120 spreads are not statistically significant. This is consistent with the fact that there is much less competition in liquidity supply for these stocks.

Some issues that arise in a fragmented marketplace are locked and crossed markets, and trade-throughs. A trade-through occurs when a trade is executed on one venue at a price that is inferior to what was available at another venue at the same time (this is, in principle, not permitted in the USA through the Regulation National Market System (NMS), although such occurrences are not prohibited in Europe where best execution is defined along more dimensions than price alone). A locked (crossed) market occurs where the bid-ask spread is zero (negative) and is evidence of poor linkages between markets and investors. In fact, such market failures are relatively infrequent and there is some evidence that they have decreased in the past few years[38], presumably through the activities of HFT arbitrageurs.

Some have argued that maker-taker pricing exacerbates liquidity issues in a fragmented marketplace. The commissioned study by Foucault (DR18) considers this question. Under this exchange-pricing structure, successful liquidity suppliers are given rebates. The stated aim is to encourage liquidity supply. Since high frequency traders are better able to place their limit orders at the top of the queue, they are the prime direct beneficiaries. Nevertheless, Foucault argues that if routing decisions are based on quotes cum fees and there is no minimum tick size, then the fee structure should be irrelevant to market outcomes. In reality these conditions do not hold: in particular, routing decisions taken by agents of traders are often not based solely on the total fee. Furthermore, tick sizes are not zero and in some cases are imposing binding constraints. He argues that maker-taker pricing in this environment may change the balance of aggressive market orders and passive limit orders and could have a beneficial effect on aggregate welfare.

Unfortunately, there is not much evidence on the effects of this pricing structure. The one empirical study of this issue, which is based on an experiment conducted on the Toronto Stock Exchange, suggests little negative impact of the maker-taker fee structure. The conclusion of review DR18 is that more empirical work is needed to allocate the costs and benefits of this market structure properly before policy options are exercised. These issues are examined in greater detail in Chapter 6 of this report.

36 O'Hara & Ye (2011).

37 DR19 (Annex D refers).

38 Ende & Lutat (2010).

In summary, most evidence points to improvement in liquidity in financial markets as a result of CBT and HFT. However, there may be some issues to do with liquidity provision in times of market stress. This general conclusion has been supported by other government-commissioned studies, including those by the Australian, Canadian and Swedish national regulators[39].

3.3.2 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[40]. He finds the total cost of trading has fallen dramatically in the period 1975 to 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 at the end of 2009. According to this study, USA large cap 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)[41] 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.

This reduction in clearing fees seems to have been replicated across European exchanges to the benefit of investors. Unfortunately, settlement fees, which are perhaps an even larger component of transaction costs, remain uncontested in Europe and have not fallen to the lower levels benefitting US investors.

Although most academic studies have shown that HFT participation has improved market quality measures such as spreads (which reflect some transaction costs), they have not directly addressed the question of whether long-term investors have been positively or negatively affected. It could be that they are targeted by high frequency traders through their speed advantage and through predatory strategies and that the main beneficiaries of the improved metrics are the high frequency traders themselves. The commissioned survey SR1 provides some support for this contention: while it showed that institutional investors generally accept that commissions and bid-ask spreads have decreased as a result of the market fragmentation and HFT facilitation of MiFID, they have concerns that most of the

39 These include:

a) The US CFTC technical advisory committee meeting http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/tac_032912_transcript.pdf Accessed: 17 September 2012.

b) The IOSCO report http://www.iosco.org/library/pubdocs/pdf/IOSCOPD354.pdf p. 26, 12: "The available evidence fails to find a consistent and significant negative effect of HFT on liquidity" Accessed: 18 September 2012.

c) Bank of Canada http://www.bankofcanada.ca/wp-content/uploads/2011/12/fsr-0611-barker.pdf p. 48. : "HFT appears to be having a profound impact on market liquidity, and its rise has coincided with an increase in trading volumes, tighter bid-offer spreads and lower market volatility, at least during periods when markets are not experiencing stress.") Accessed: 17 September 2012.

d) Swedish FSA http://www.mondovisione.com/media-and-resources/news/swedish-financial-supervisory-authority-finansinspektionen-high frequency-tra/, " investigation has demonstrated that the impact of HFT on trading is smaller than feared. Swedish investors believe that trading has undergone a transformation and that the market has become more volatile, but that these changes can be explained by multiple factors and not only the emergence of HFT" Accessed: 18 September 2012.

e) Australian Regulators http://www.asic.gov.au/asic/pdflib.nsf/LookupByFileName/rep-215.pdf/$file/rep-215.pdf, p. 91: "The benefits to the economy and the equities industry from competition are likely to outweigh the costs of order flow fragmentation, which include those of increased surveillance, technology and information. The net benefit will be positive if competition is introduced with the proposed rules framework"; p. 94: "Some current industry players will thrive in the new conditions, while others will find the environment extremely challenging. Those with the technology and volume to gain full benefit from the multiplicity of execution venues may experience lower costs and new business/trading opportunities" Accessed: 17 September 2012

f) The LSEG submission on MiFID http://www.londonstockexchange.com/about-the-exchange/regulatory/lseg-submission-to-ec-on-mifid02-02-11.pdf : Regarding AT/HFT, p. 1: "No evidence of market failure" Accessed: 17 September 2012.

40 Jones (2002).

41 DR16 (Annex D refers).

benefits are captured by their brokers. These investors believe that infrastructure costs have increased as well as some other indirect costs[42].

The interests of institutional investors are of great importance. A commissioned study by Brogaard *et al*. (DR21) examines the direct effects of HFT on the execution costs of long-term investors. The authors use a new UK dataset obtained from the detailed transaction reports of the UK Financial Services Authority (FSA) over the period 2007 to 2011 to provide a better measurement of HFT activity. They combine this with Ancerno data on institutional investors' trading costs. To test whether HFT has impacted the execution costs of institutional traders, the authors conduct a series of event studies around changes in network speeds on the LSE to isolate sudden increases in HFT activity. This study found that the increases in HFT activity have no measurable effect on institutional execution costs. Of course, additional studies linking HFT and institutional trading costs in other market settings would be helpful in determining the generality of this finding.

In summary, the evidence is that transaction costs have declined in the past decade, mostly due to changes in trading market structure (which are closely related to the development of HFT). Specifically, there has been increased competition between trading venues and greater use of labour-saving technology.

3.3.3 Price efficiency

A central claim of financial economists is that more efficient prices (which better reflect fundamental values) in financial markets contribute to more informed financing and investment decisions and, ultimately, to better allocation of resources in the wider economy and, hence, higher welfare. The holy grail of fully efficient markets is now widely acknowledged to be impossible, following the work of Grossman & Stiglitz (1980)[43]: if obtaining information and trading on it is costly, then no one will gather it unless they can profit, but if all the information is already in prices, then no one can profit from it. If no one gathers information, then how can it be impounded in prices? In the light of this, the relevant question is how inefficient markets are and with respect to which information.

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. The usual method of measuring the degree of market inefficiency is through the predictability of prices based on past price information alone. In practice, widely used measures such as variance ratios and autocorrelation coefficients estimate the predictability of prices based on linear rules.

Several studies commissioned by this Project address the question of whether HFT strategies are likely to improve or worsen price efficiency. Brogaard (DR10) describes how high frequency traders 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 liquidity supply and statistical pattern detection used in so-called 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. Farmer (DR6) cautions that, as market ecologies change, the transition to greater efficiency may be slow.

42 One reason for increased transaction costs for large traders may be quote matching. Suppose that a large trader places a limit order to buy at 30. A clever trader who sees this order could immediately try to buy ahead of it, perhaps by placing an order at 30 at another exchange, or by placing an order at a tick better at the same exchange. If the clever trader's order fills, the clever trader will have a valuable position in the market. If prices subsequently rise, the trader will profit to the extent of the rise. But if values appear to be falling, perhaps because the prices of correlated stocks or indices are falling, the clever trader will try to sell to the large trader at 30. If the clever trader can trade faster than the large trader can revise or cancel his order, and faster than can other traders competing to fill the large trader's order, then the clever trader can limit his losses. The clever trader thus profits if prices rise, but loses little otherwise. The large trader has the opposite position: if prices rise, he may fail to trade and wish that he had. If prices fall, he may trade and wish that he had not. The profits that the clever trader makes are lost profit opportunities to the large trader. The quote-matching strategy was established long before HFT, however. See Harris (1997) *Order Exposure and Parasitic Traders*.

43 Grossman & Stiglitz (1980).

Brogaard *et al.* (2012)[44] find that high frequency traders play a positive role in price efficiency by trading in the direction of permanent price changes and in the opposite direction of transitory pricing errors on average days and the days of highest volatility. This is done through their marketable orders. In contrast, HFT passive orders (which are limit orders that are not immediately marketable) are adversely selected in terms of the permanent and transitory components as these trades are in the direction opposite to permanent price changes and in the same direction as transitory pricing errors. HFT marketable orders' informational advantage is sufficient to overcome the bid-ask spread and trading fees to generate positive trading revenues. Non-marketable limit orders also result in positive revenues as the costs associated with adverse selection are smaller than the bid-ask spread and liquidity rebates.

Negative effects on efficiency can arise if high frequency 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 volume weighted average price (VWAP) 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, such as 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 with the growth of CBT. Castura *et al.* (2010) investigate trends in market efficiency in Russell 1000/2000 stocks over the period 1 January 2006 to 31 December 2009[45]. Based on evidence from intraday[46] variance ratios, they argue that markets become more efficient in the presence of and increasing penetration by HFT.

Johnson and Zhao (DR27) look at ultra-high frequency data and find some evidence that (relatively) large price movements occur very frequently at the machine time-scale of less than half a second. Brogaard (2010) provides further evidence on this issue. He estimates that the 26 HFT firms in his sample earn approximately $3 billion in profits annually. Were high frequency traders not part of the market, he estimates that 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 by an additional $.056. He also shows that HFT trades and quotes contribute more to price discovery than does non-HFT activity.

Linton (DR1) provides evidence based on daily UK equity data (FTSE All-Share). Specifically, he computes variance ratio tests and measures of linear predictability for each year from 2000 to 2010. The measures of predictability (inefficiency) fluctuate around zero with sometimes more and sometimes less statistical significance. During the financial crisis there was somewhat more inefficiency, but this declined (until late 2011 when there was a series of bad events in the market). Overall, DR1 found no trend in daily market efficiency in the UK market, whether good or bad. Figure 3.3 shows an updated graph showing the variance ratio of the FTSE All-Share index until June 2012[47].

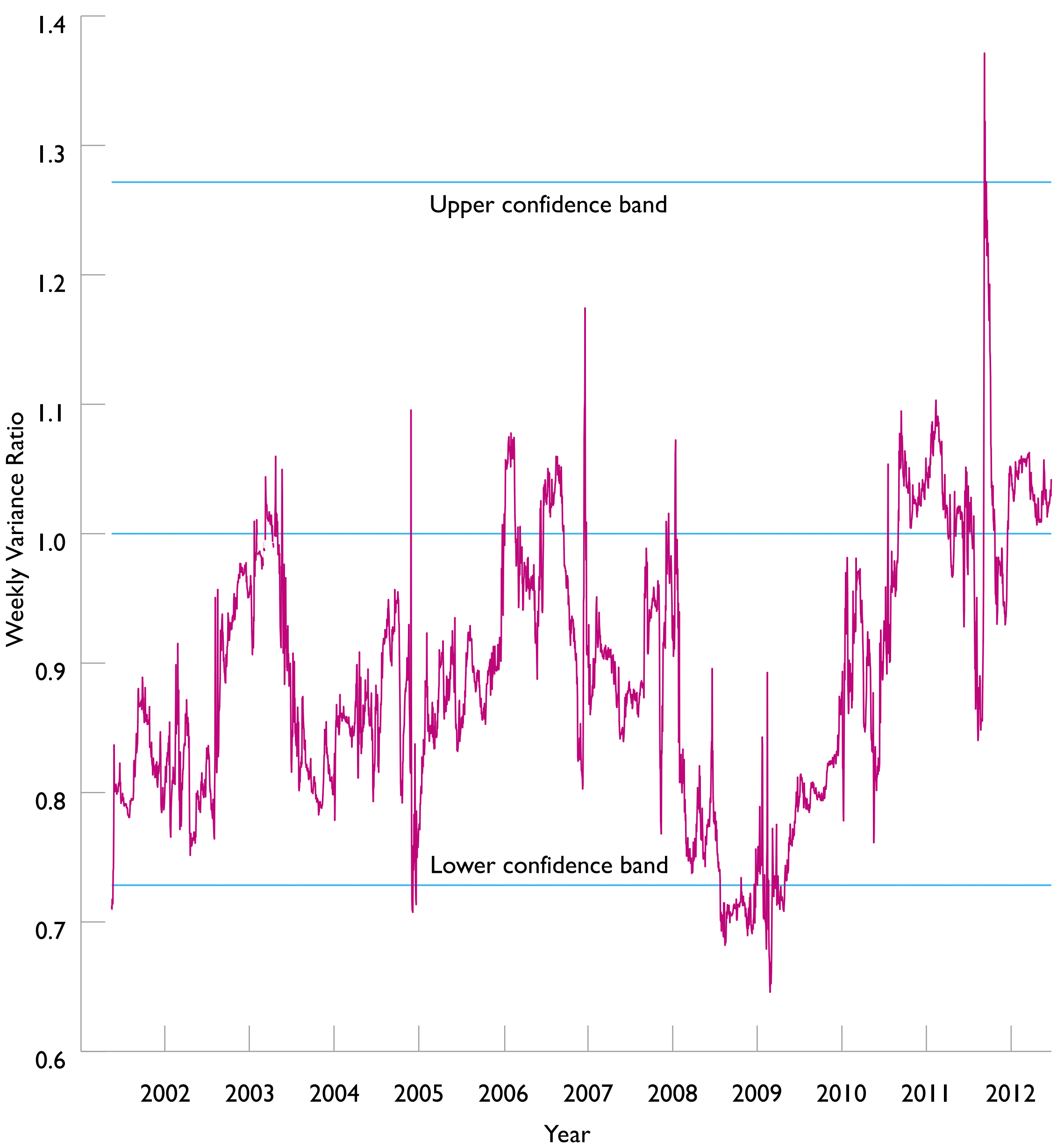
44 Brogaard *et al.* (2012).

45 Castura *et al.* (2010), *op. cit.*

46 They look at 10:1 second ratios as well as 60:10 and 600:60 second ratios.

47 This is calculated using a rolling window of 250 daily observations (roughly one trading year). It is the weekly variance ratio which is the variance of weekly returns divided by five times the variance of daily returns. If returns were unpredictable (i.e., markets were efficient) this ratio should be one, which is displayed by the middle dashed line. The upper and lower dashed lines represent 95% confidence intervals centred at one. There are few instances when the confidence bands over this period are exceeded, with notable exceptions in 2008/2009 and a brief violation in 2011. The daily return series is of interest to many retail investors.

Figure 3.3: FTSE All-Share daily return predictability 2000-2012



Source: FTSE

One further issue is the location of price discovery and whether this has been affected by the increased competition between venues. This is an important issue in the European context where the few regulated markets (like the LSE) claim a unique position, based partly on their supposed pivotal role in price discovery. Gresse (DR19) finds that active multilateral trading facilities (MTF), such as Chi-X in London, played a significant part in price discovery. Furthermore, this transfer of price discovery has not worsened the efficiency of the consolidated market in Europe.

These studies have all addressed the functioning of the secondary market for equity trading. Boehmer *et al.* (DR23) have investigated the effects of AT on firms' equity capital-raising activities in an international panel data-set. They find some evidence that the yearly AT intensity (which they measure by the message traffic in a stock) reduces net equity issues over the next year with the main channel being the effect on share repurchases. It is not clear what the mechanism is for this effect, so these statistical results should be treated with some caution.

In summary, the preponderance of evidence suggests that HFT has not harmed, and may have improved, price efficiency. It may also follow from this that HFT has improved welfare, although the connection between the two concepts is complex[48].

3.4 What has been the impact of computer-based trading on market quality?

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 Commodity Futures Trading Commission (CFTC) and SEC staff. These reports show that a complex interaction of forces led to the Dow Jones Industrial Average falling 998 points, the largest intraday drop in US market history. While the Flash Crash lasted less than 30 minutes, for a brief interval more than one trillion dollars in market capitalisation was lost. In the aftermath of the crash, more than 20,000 trades were cancelled. A more lasting effect has been a sustained withdrawal from equity investing by retail traders.

The CFTC and 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, internalisers (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 CBT 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 AT strategies are triggered in response to automatic data feeds and so have the potential to lead to cascades in market prices as selling causes price falls which prompt 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 underlines that market illiquidity is not a new event. What matters for this Project is whether current computer-driven practices are causing greater illiquidity risk. Madhavan (2011) argues that fragmentation linked to HFT may be one such cause[49]. He shows that fragmentation measured using quote changes, which he argues is reflective of HFT 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[50]. They found that high frequency traders initially acted as liquidity providers but that, as prices crashed, some high frequency traders 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 *et al.* (2011) argue that historically high levels of order toxicity forced market makers to withdraw during the Flash Crash[51]. 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.* (2011a) 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.

48 Jovanovic & Menkveld (2011) provide a model that formalises the role that high frequency traders perform as middlemen between 'real' buyers and 'real' sellers. They compare a market with middlemen to a market without. Their model allows for both positive and negative effects (of middlemen) depending on the value of certain parameters that measure the informational advantage that high frequency traders would have through their ability to process numerical information about the order book rapidly. They calibrate their model using the empirical work presented in Menkveld (2012) about the entry of an HFT market maker on Dutch index stocks on Chi-X. Empirically, HFT entry coincided with a 23% drop in adverse-selection cost on price quotes and a 17% increase in trade frequency. They conclude that, overall, there is a modest increase in welfare.

49 Madhavan (2011).

50 Kirilenko *et al.* (2011).

51 Easley *et al.* (2011a).

Other factors were at play during the Flash Crash; for example, the delayed and inaccurate data feeds and the breaking of the national market system through self-help declarations by various exchanges[52].

There have been a variety of other, smaller illiquidity events in markets since the Flash Crash. On 8 June 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 2,000–3,000 orders in a single second, causing an eightfold spike in volatility and moving the oil price $1 before the algorithm was shut down. In March, trades in ten new Morningstar exchange traded funds (ETF) 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 error made while entering data). On 24 August 2010, there were very rapid changes in the prices of five LSE-listed stocks: BT Group PLC, Hays PLC, Next PLC, Northumbrian Water Group PLC and United Utilities Group PLC. In this case, there appear to have been no significant news about macro-economic or stock-specific fundamentals. The rapid falls triggered the LSE circuit breakers, which effectively prevented further falls and contagion across the broader market. Chapter 6 discusses the efficacy of the existing market controls.

There are also many cases where prices have moved substantially within a day driven by fundamental news rather than by trading system issues. The 1929 stock market crash is widely believed to be a correction of previous overvaluations, and the 1987 crash is also believed to be at least partly driven by fundamentals, since it was echoed around the world. Many firm-specific crashes have occurred that were due simply to bad news about the performance prospects of the companies involved[53].

In summary, the evidence suggests that HFT and AT may be contributing to periodic illiquidity in current markets. They may not be the direct cause of these market crises, but their trading methods mean that their actions or inactions may temporarily amplify some liquidity issues.

3.5 Future: how is the impact of computer-based trading on liquidity likely to evolve in the next ten years?

There is considerable uncertainty regarding the future role of HFT. TABB Group (2012) estimates that HFT accounted for 53% of trading in US markets in the first half of 2011, a decrease from the 61% share it held in 2010[54]. 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 HFT made up 75% or more of US volume between 4 and 10 August 2011[55]. Whether HFT will continue to be as active when volatility subsides to normal levels remains to be seen[56].

There are also indications that the profitability of HFT is reaching its limits and in the next ten years may come under further pressure. Such reductions may arise for a variety of reasons: the potential move to sub-penny pricing in the USA 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 make their business model unviable. 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.

52 See http://www.sec.gov/divisions/marketreg/rule611faq.pdf for an explanation of the self help exception the NMS rules. Essentially, exchanges are allowed in exceptional circumstances (such as those occurring on May 6th, 2010) to not exchange messages with other exchanges, which would be their obligation under Regulation NMS. Accessed: 17 September 2012.

53 For a recent example, Thomas Cook (a FTSE 250 company at the time) closed on 21 November 2011 at 41.11 and closed at 10.2 on 22 November 2011. The 75% loss of value was attributed to an announcement of bad results that took place before trading began. Most of the value loss was present at the opening of the market on the 22 November but the decline continued rapidly through the rest of the day (to the extent that it triggered a number of trading halts). There is some perception that this loss of value took place at a much faster pace than it would have done ten years ago.

54 TABB Group (2012).

55 Mehta (2011).

56 The increased activity of HFT during this volatile period seems at odds with some earlier arguments that their liquidity provision can reduce during times of market stress. Perhaps this illustrates the difference between regular volatility, which has both upside and downside price movements, and a market fall where only downside price changes occur.

Nonetheless, it seems inevitable that CBT will remain a dominant force in markets over the next ten years. One reason for this will be technological advances that facilitate the automated extraction, aggregation and filtering of news[57]. Such news analytics, discussed in Chapter 2, could be used in model construction for high frequency traders, as well as for portfolio managers. News analytics 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 CBT strategies provides a means for traders to capitalise on information before it is incorporated into market prices. High frequency traders will be well positioned to take advantage of such nascent technology.

To the extent that 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 advanced traders. Consequently, it may be that all trading evolves towards CBT, reflecting the fact that technology diffuses across populations given time.

As this happens, market systems may experience unwanted negative effects. 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 many 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 6 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 USA. When this occurs, market liquidity is affected[58].

A related systematic risk can arise if a large sub-set of market participants are following the same strategies. For example, if news analytics 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 to be able to function[59]. The future of CBT may thus involve more technology capable of controlling the technology controlling the markets.

Finally, it is not possible to legislate market crashes away, any more than it is possible to dispense with the business cycle. In future crises, regulators need to be able to reconstruct what has happened; when, where and why. One of the consequences of MiFID was the loosening of the trade reporting protocols in Europe so that trades can be reported through Trade Reporting Facilities (TRF) other than the 'primary' market. This made *ex-post* auditing work much more difficult than it is in the USA where they have had a consolidated tape for some time. Europe is somewhat behind in the collection, standardisation and analysis of financial data in comparison to the USA, where the Office of Financial Research (OFR) has been commissioned by the Dodd-Frank Act to found a financial data centre to collect, standardise and analyse such data[60].

57 See DR8 (Annex D refers).

58 At the time of writing, many exchanges, including the LSE, have introduced 'message throttling' as well as pricing systems which may mitigate some of these effects.

59 Easley *et al.* (2011b).

60 See the OFR Annual Report (2012).

3.6 Conclusions

CBT 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. The scale of improvements may be fairly small and, in the short term, they may have been obscured by the background of a very poor performance by Organisation for Economic Co-operation and Development (OECD) economies and stock market indexes in particular. However, 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 realities of trading in asset markets. Caution must be exercised to avoid undoing the 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.

The next chapter will consider how the increased use of technology in markets affects financial stability.


Cur P/E Level
Est P/E Level
Price USD
High
Low
Average

4 Financial stability and computer-based trading

Key findings

Economic research to date provides no direct unambiguous 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: 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 or multiple algorithms, 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 (for example, extremely rapid crashes), until a disastrous failure occurs.

4 Financial stability and computer-based trading

4.1 Introduction

As described in Chapter 1[1] a broad interpretation of computer-based trading (CBT) is used in this report. A useful taxonomy of CBT identifies four characteristics which 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. Third, they may be classified as engaging in either uninformed or informed trading and fourth, a CBT algorithm generates the trading strategy or only implements a decision taken by another market participant.

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' use of CBT cannot be neatly assigned to only one of the above four categories, 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 named financial or trading corporation, such as a specific investment bank or fund management company. For much the same reasons, the focus of the discussion in this chapter is not on any one asset class (such as equities, foreign exchange (FX), commodities or government bonds), but rather on the forces that seem likely to shape future issues of stability arising from CBT. This chapter summarises 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. There is no agreed definition of 'systemic stability' and 'systemic risk', and the reader is referred to DR29[2] for a discussion and a survey on empirical systemic risk measures.

4.2 How has computer-based trading 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[3] by enabling a timely and reliable reaping of mutual gains from trade, as well as allowing investors to diversify their holdings. 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 can lead also to private and social costs. The focus of this chapter is solely on possible repercussions of CBT (including high frequency trading (HFT) in particular) 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. It is hoped 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 conclusions of this chapter apply to any given market structure, but they are especially relevant to the continuous double auctions of the electronic limit order book 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. Even in a huge market such as the FX market, a sufficiently large order

1 See Box 1.1 in Chapter 1, and also DR5 (Annex D refers).

2 DR29 (Annex D refers).

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

can temporarily sway prices, depending on how many other orders are in the market (the 'depth' of the market) at the time and how quickly the book is replenished (the 'resilience' of the market).

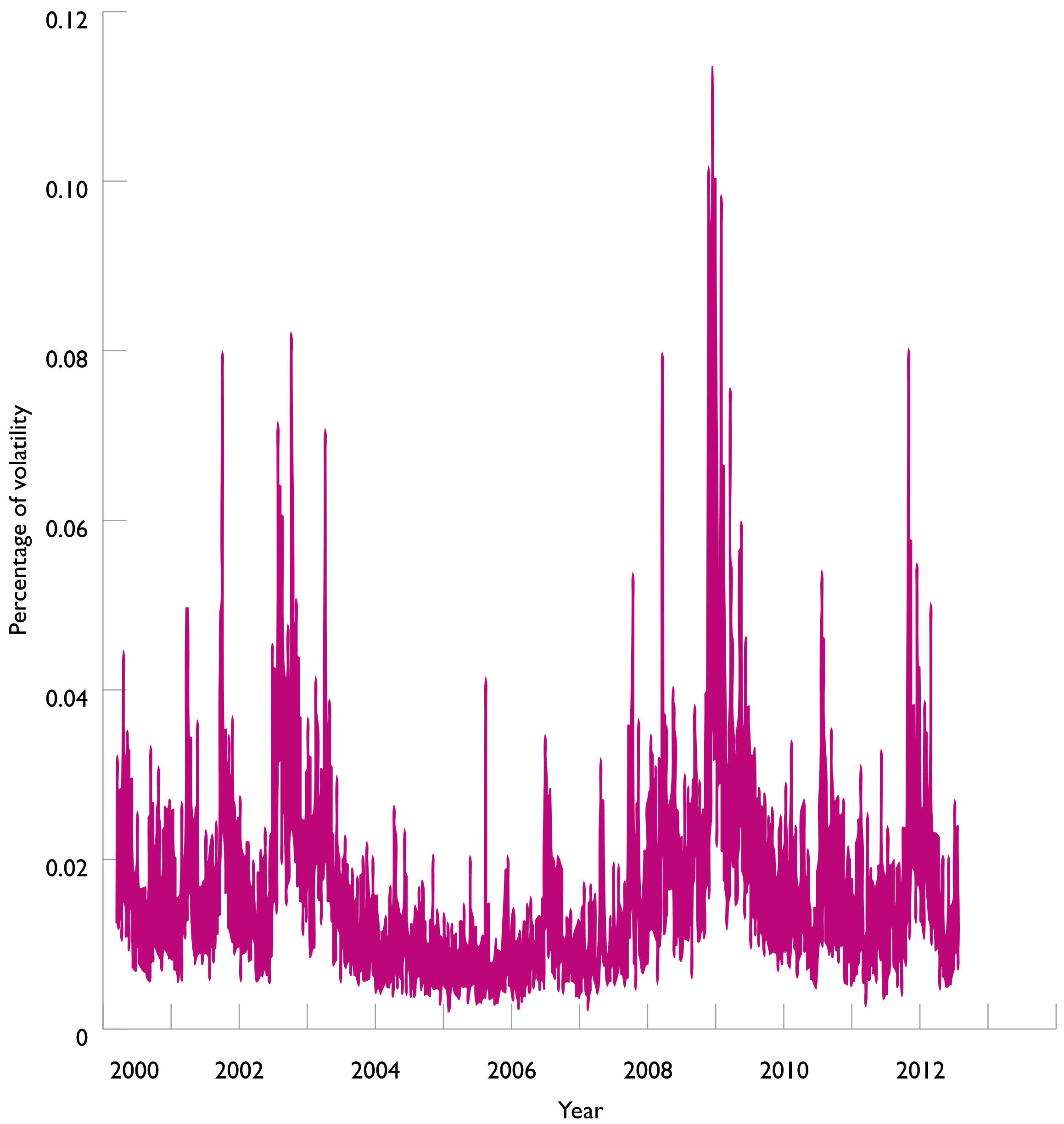
Price volatility is a fundamental measure useful in characterising financial stability, since wildly volatile prices are a possible indicator of instabilities in the market and may discourage liquidity provision[4]. In DR1, Linton notes that fundamental volatility has decreased in the UK equities market since the turmoil of 2008/2009, and liquidity and trading volume have slowly returned[5]. If HFT contributes to volatility, Linton argues, it might be expected that the ratio of intraday volatility to overnight volatility would have increased as HFT became more commonplace, but they find no evidence to support that hypothesis. They note that the frequency of large intraday price moves was high during the crisis period, but the frequency has declined to more normal levels since the end of 2009. However, Boehmer *et al.* (2012)[6] in a study spanning 39 exchanges in 36 countries have found that higher volatility and CBT activity move together over the period 2001 to 2009, though causality is not yet clear and the economic magnitudes appear to be small.

4 Stability can, of course, differ from volatility by placing significantly greater weight on large, infrequent price changes, especially if the latter do not appear to be fundamental.

5 See Figure 4.1 for a time series of realised volatility computed as (high-low)/low; the implied volatility index VFTSE follows a similar pattern.

6 Boehmer *et al.* (2012).

Figure 4.1: FTSE100 volatility between 2000-2012



Source: FTSE

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[7], not least because it is not clear whether HFT increases volatility or whether volatility invites more HFT – or both or neither[8]. Significant challenges in evaluating HFT arise because much of its growth coincides with the 2008/2009 turmoil and data fully characterising HFT are lacking[9]. Indirect studies of CBT and HFT provide interesting evidence highlighting the importance of further study with better data, but are subject to various interpretations[10].

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 of these could, in principle, lead to system-wide failure events.

However, the more interesting and significant aspects of analysing financial stability relate to the general *nonlinear dynamics* of the financial system. Put simply, a nonlinear system is one in which a given change in one variable may lead either to a small or a large change in another variable, depending on the 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 long-term behaviour because an observer can never know the relevant key values with sufficient accuracy. Furthermore, because economic systems are composed of agents (such as individuals, firms and regulators) that interact in various ways, the dynamics of the system can depend on precisely how these agents interact[11]. The scientific literature on the complex, nonlinear dynamics of networked systems is currently in its infancy in terms of concrete predictions and reliably generalisable statements. But, integrating existing knowledge with the foundations of financial economics literature can offer glimpses of important insights. Figures 4.2 to 4.6 illustrate unusual price, quote and volume dynamics generated by algorithms consistent with feedback loops, both of the amplifying (positive feedback loop) type and of the pinning (negative feedback loop) type.

7 See DR1, DR 12 (Annex D refers), Brogaard, (2010), Chaboud *et al.* (2009), and Hasbrouck & Saar (2011).

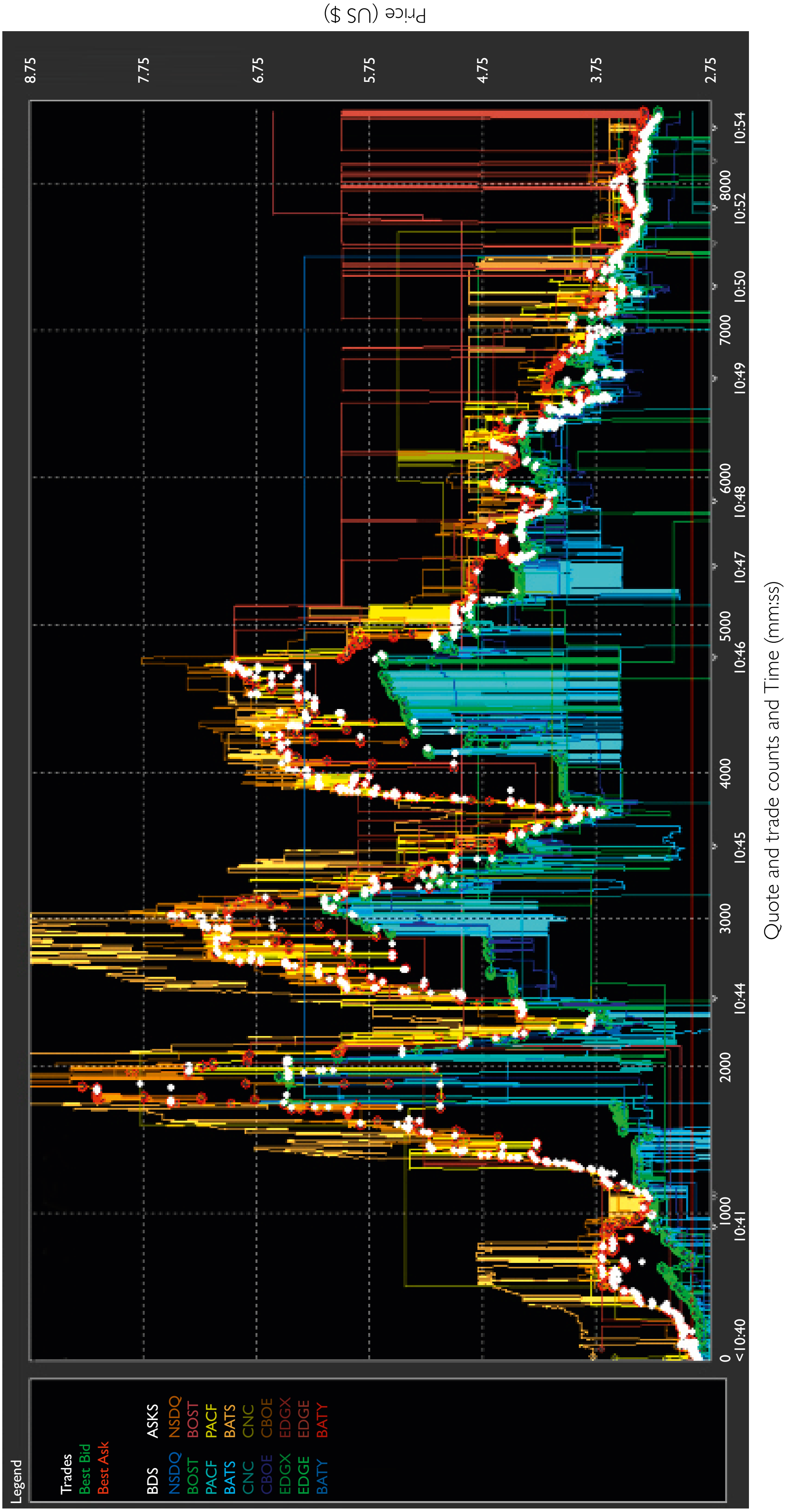
8 It is well-known that volatility alone accurately represents neither risk nor welfare (see, for example, Rothschild & Stiglitz (1970) and Grossman, (1976)). Therefore, without further studies, no automatic and unambiguous welfare conclusions can be reached from observations on volatility.

9 Jovanovic & Menkveld (2011), have a comparison of the volatility of Dutch and Belgian stocks before and after the entry of one HFT firm in the Dutch market and find that the relative volatility of Dutch stocks declines slightly.

10 For example, Zhang (2010), 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. A stronger relation between volume and volatility can result from increased welfare-enhancing investor risk-sharing. Therefore, indirect studies of HFT and CBT as in Zhang (2010) are not recommended as a basis for formulating policy options.

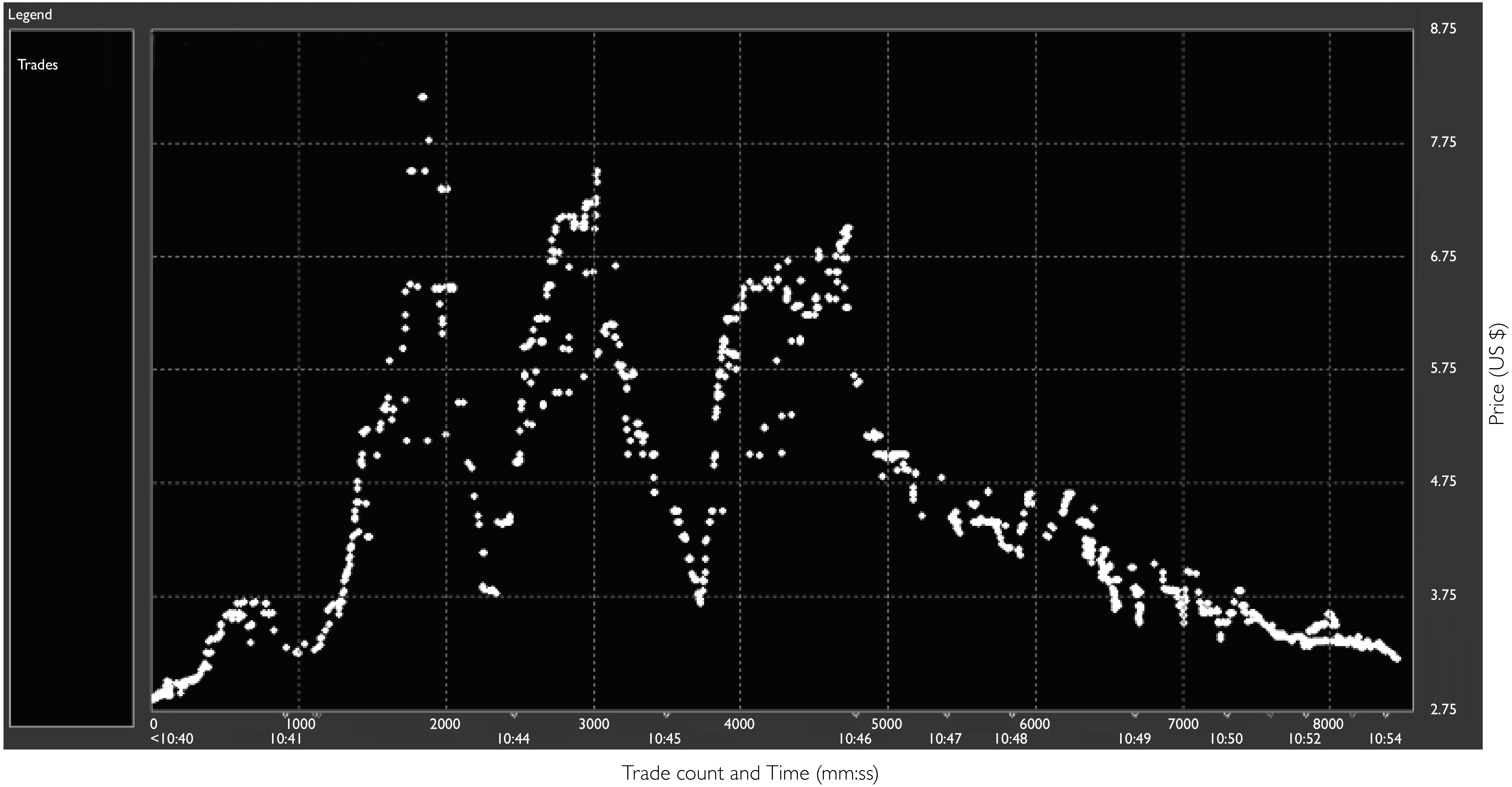
11 See DR6, DR7 (Annex D refers), and Haldane & May (2011).

Figure 4.2: Violent cycling of a stock with ticker CNTY in different USA exchanges on 21 June 2011



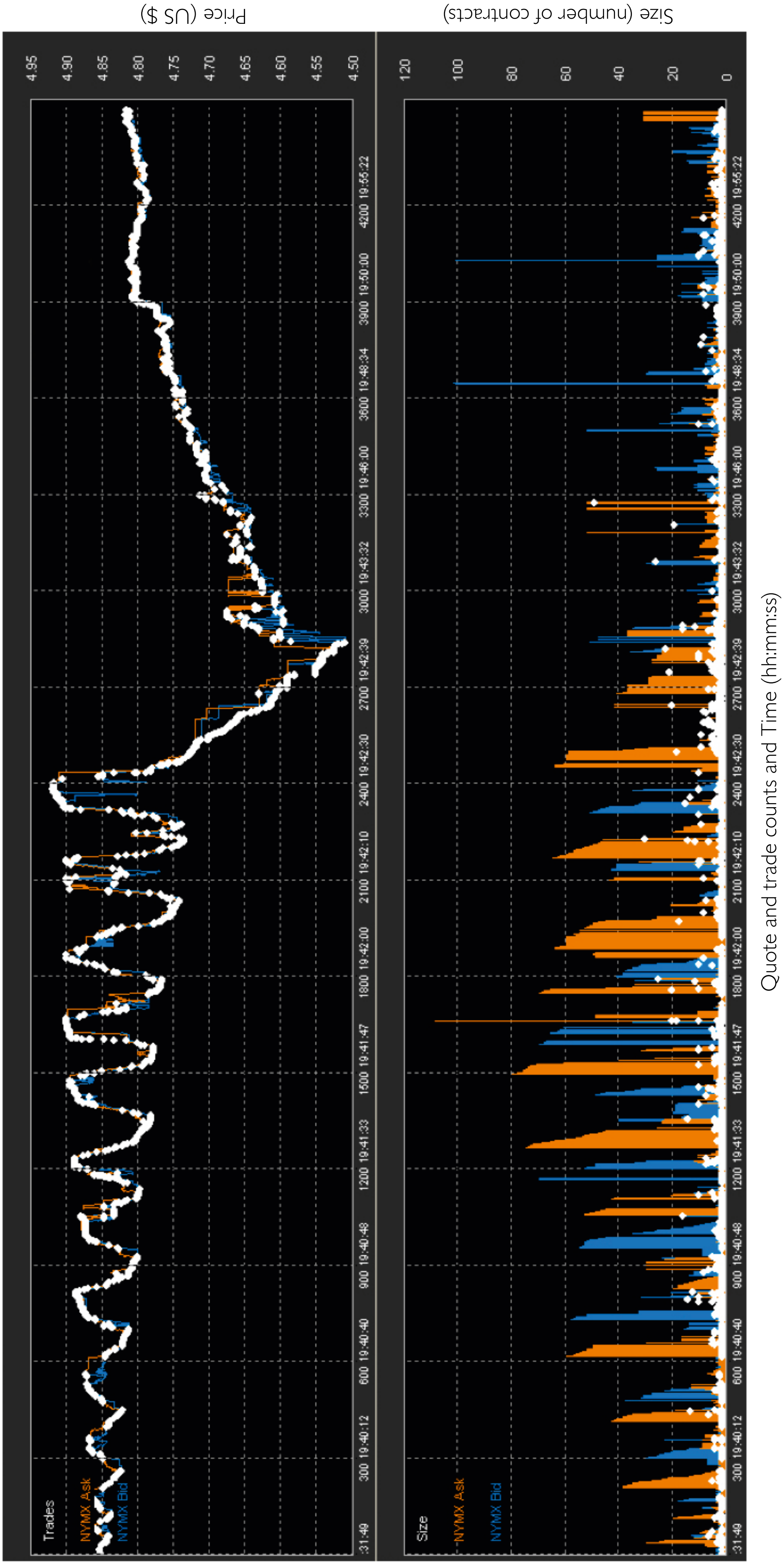
Source: Nanex

Figure 4.2 (cont.): Violent cycling of a stock with ticker CNTY in different USA exchanges on 21 June 2011



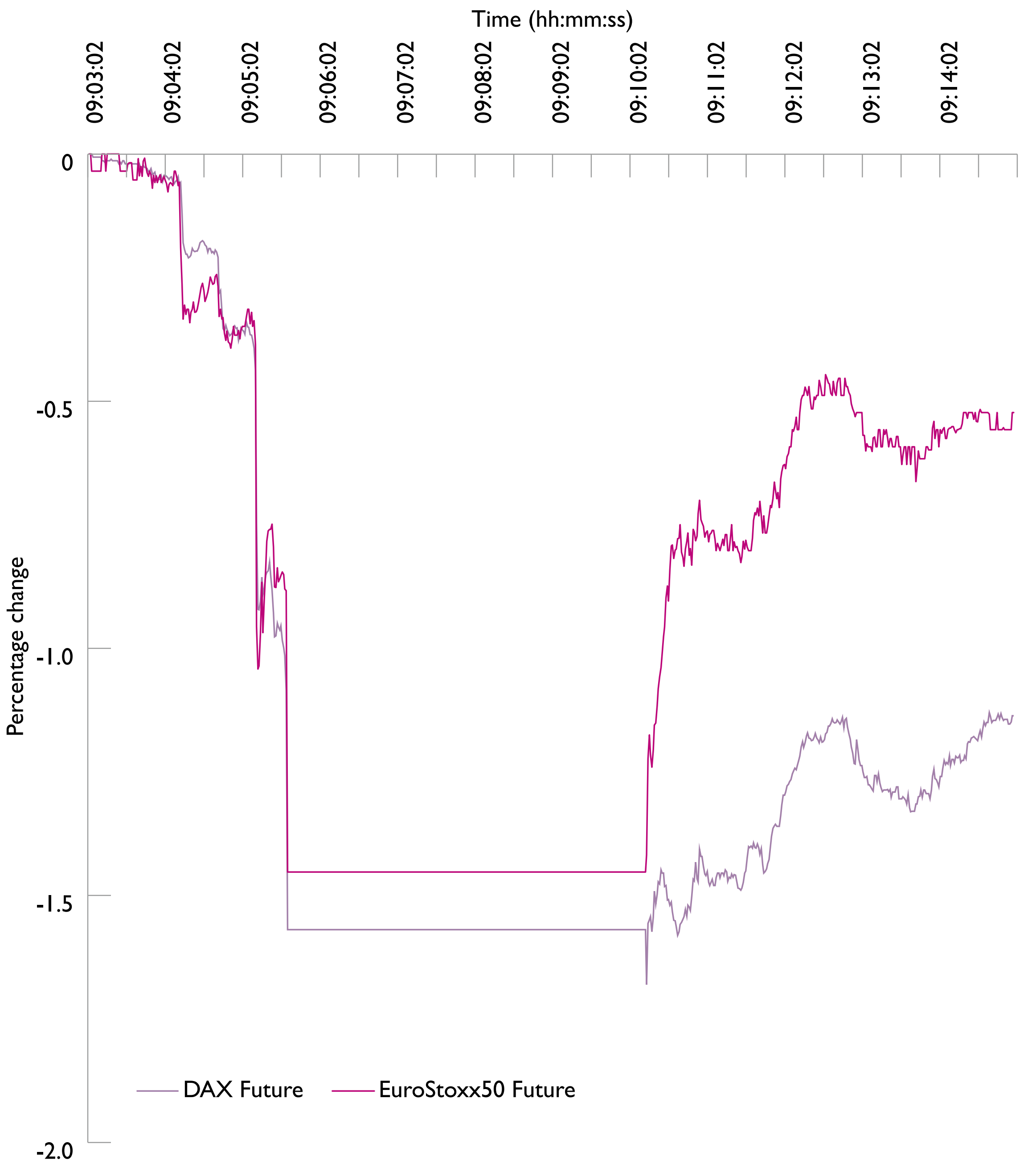
Source: Nanex

Figure 4.3: Violent cycling of natural gas futures prices on NYMEX, 8 June 2011 with a resulting crash and subsequent recovery



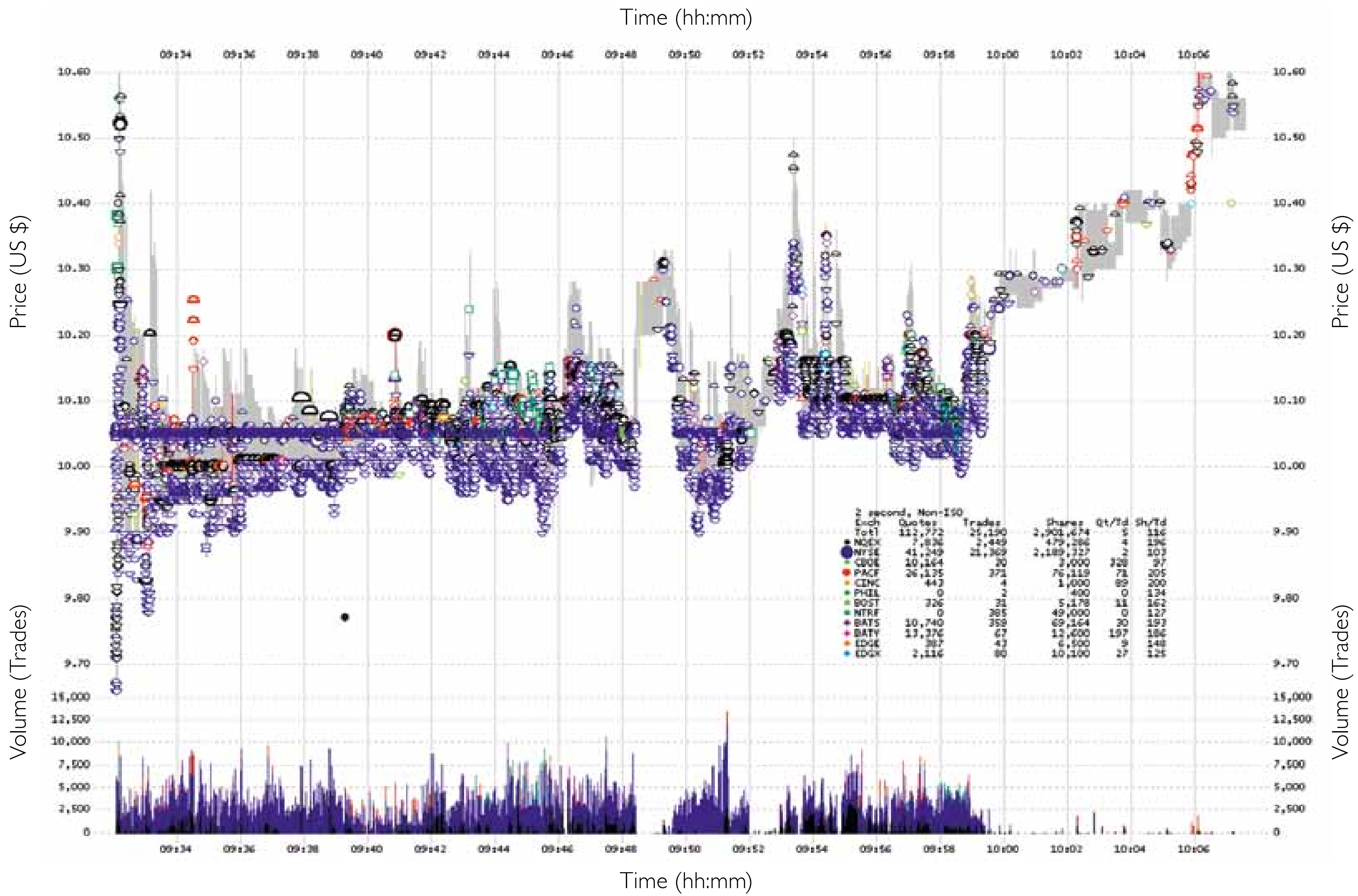
Source: Nanex

Figure 4.4: An unexplained slump in European index futures on 27 December 2010



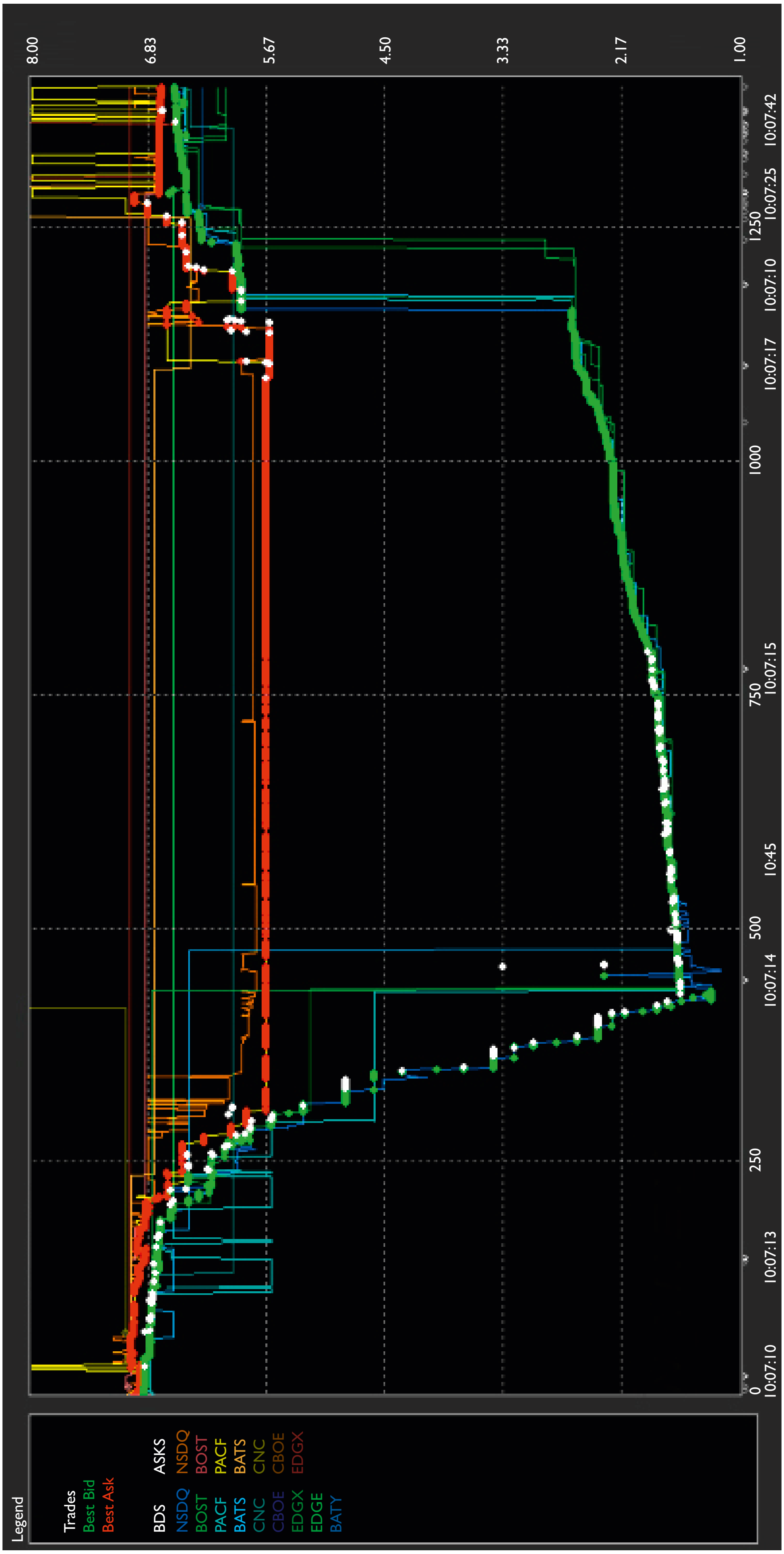
Source: Deutsche Börse Group: Eurex Exchange

Figure 4.5: A pinning example that occurred during the recent Knight Capital algorithm malfunctioning episode where algorithms bought and sold in a way that kept prices of a stock with ticker CODE glued (pinned) around $10.05



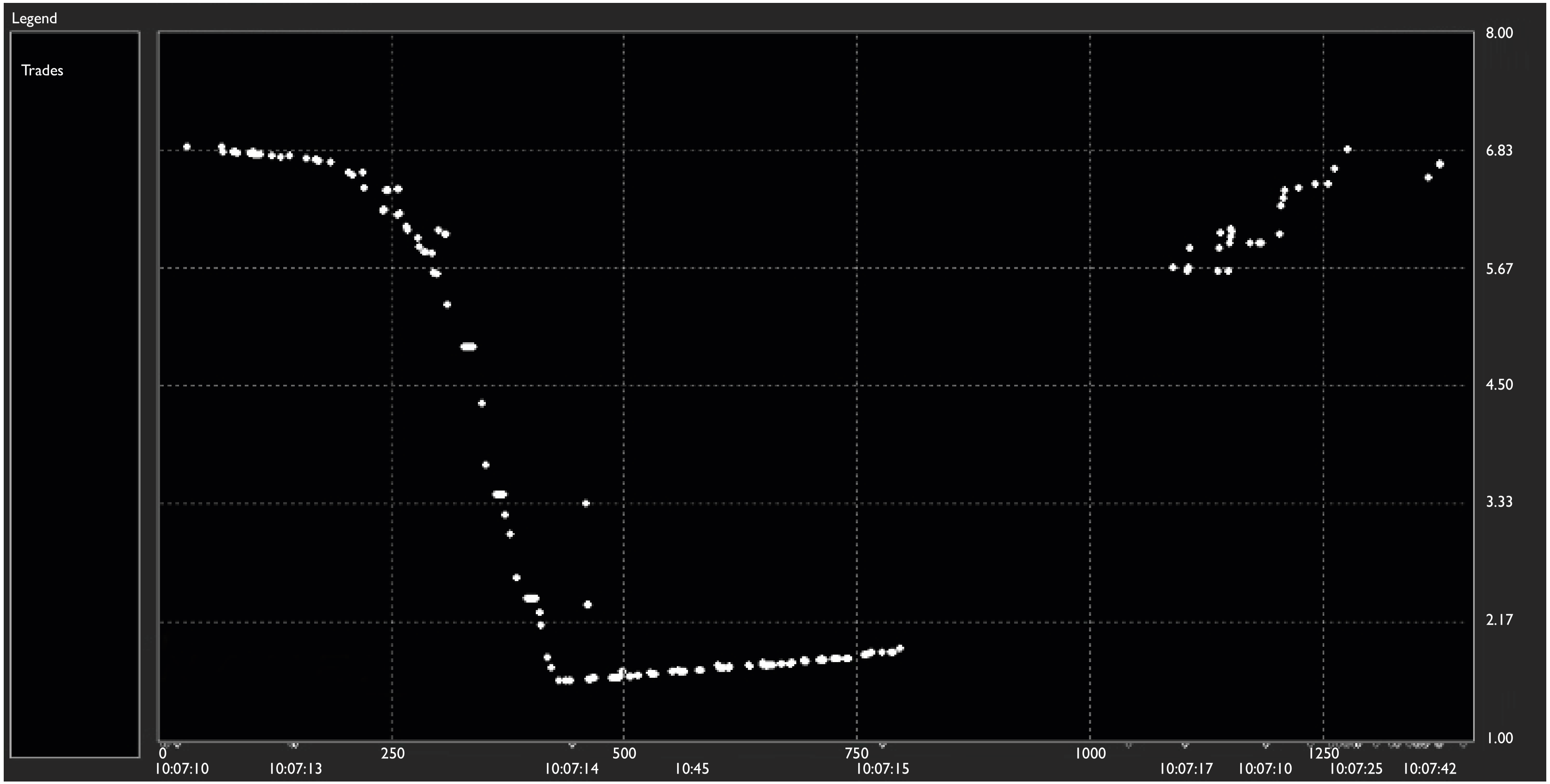
Source: Nanex

Figure 4.6: A crash in the stock AMBO that occurred and resolved itself within a single second



Source: Nanex

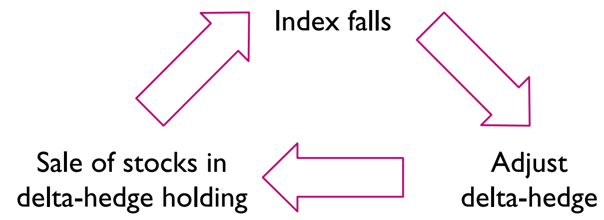
Figure 4.6 (cont.): A crash in the stock AMBO that occurred and resolved itself within a single second



Source: Nanex

While market crashes have always been a feature of financial markets, the problem of understanding the mechanisms behind system-level events in CBT environments is more recent. A good illustration for the sort of systemic non-linear events that mechanical rule-following is able to generate can be found in the portfolio-insurance-led market decline of 1987[12]. In order to hedge their risks, when stock indices dropped, portfolio insurers were required to adjust their 'delta-hedge' holdings 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, when the index dropped, stocks held in the delta-edge portfolio were sold, which depressed prices and pushed the index even lower; this then caused further sales of the delta-hedge holdings, pushing the index still lower. This positive feedback loop (the effects of a small change looping back on itself 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 4.7 where a fall in the value of an index forces delta-hedgers to sell into a falling market, which in turn puts downward pressure on the value of the index. 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 'programme trades' led in 1987 to huge selling pressure and to price falls which were much deeper than were warranted by actual market conditions.

Figure 4.7: Hedge feedback loop



Forced sales in a nonlinear, self-fulfilling and sometimes self-exciting frenzy create a discrepancy between market prices and the true values of securities (measured by the fundamental value of the asset) and thus constitute 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 is of such wild feedback loops in a financial system. This endogenous risk is the logical thread that runs through much of the rest of this chapter. The endogenous risk from human-programmed algorithms may differ in important ways from feedback loops and risks in markets with greater direct human involvement.

4.3 How is computer-based trading thought to affect financial stability?

4.3.1 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 has the potential to 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. Some of the insights into what the precise mechanisms could be are outlined below.

Three key inter-related 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[13]. The mechanisms of deviation may be the information sensitivity mechanisms, and/or the internal feedback loops, both of which are discussed

12 See DR9 (Annex D refers), and also Gennotte & Leland (1990).

13 See DR7 (Annex D refers), and also Haldane & May (2011), *op. cit.*

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. At very high speeds, sensitivity may also be positively related to speed.

- **Information:** the existence of excessive nonlinear sensitivities can be due to informational issues: in other words, who knows what and when. The information structure of a market has the potential to increase 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[14]. In these game scenarios, 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 whether creditors run or stay. With diffuse information, violent cascades of failures over an entire market system can be triggered by such small events[15].
- **Endogenous risk:** this term, commonplace in the financial literature[16], 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[17].

In the section that follows, a number of feedback loops that can contribute to endogenous risk are explored. One of the reasons the sensitivities and feedback loops at very high speeds may look qualitatively different from those at lower speeds, even if accelerated, lies in the fact that beyond the limits of human response times, any feedback loops must be generated solely by robots[18]. It is plausible that the properties of securities return time series at different frequencies are very similar to each other through scaling[19], at least as long as the frequency is low enough so that micro structure effects do not come into the picture. This similarity may no longer hold beyond a critical frequency at which only machines can react to each other. One of the reasons may be the discontinuity in the set of possible feedback loops as one goes to very small scales and excludes human interactions[20]. While 'mini-flash crashes' occur regularly at very small scales, they are mostly self-healing in a matter of milliseconds and aggregate to noise without attracting much attention[21]. This does not mean that a mini-flash crash of the feedback loop variety (as opposed to an execution algorithm simply splitting a large trade and continuously buying or selling) could not form a chain reaction, gather momentum and be visible at human scales also, or indeed lead human traders to feed into the loop.

14 See, for example, Carlsson & Van Damme (1993) and Morris & Shin (2002).

15 Information, sensitivity and speed interact. For instance, with lower tick sizes the frequency of order book updating increases dramatically. Given possible and realistic latencies for traders, there would be inevitable information loss: no trader would at any moment in time be able to have an accurate idea of the current state of the order book, let alone an accurate depiction of multiple order books across trading venues. Very fast markets make concurrent precise order book information virtually impossible and lead to trading decisions taken under a veil of incomplete knowledge, not only about the future but also about the present. Trading strategies need to second-guess the state, possibly leading to over-reactions (see Zigrand (2005), for a formal model). In that sense, speed influences both information and sensitivity according to a U shape: perhaps, as with a bicycle, both too low a speed and too high a speed lead to inefficiencies and instabilities.

16 See Danielsson & Shin (2003), Danielsson *et al.* (2011) and Shin (2010).

17 See DR2, DR6, DR7, and DR9 (Annex D refers).

18 See DR6, DR7 and DR27 (Annex D refers). DR27 suggests a transition may occur at around 1000 microseconds.

19 This is the so-called characteristic of self-similarity, roughly saying for instance that the distribution of hourly returns is very similar to the distribution of minute by minute returns when minute by minute returns are multiplied by the appropriate scaling factor of 60 to a power H. For example in the Gaussian case H=1/2. In that sense, self-similarity means that short-dated returns really are just a sped-up version of lower frequency returns, with no qualitative break in their statistical behaviour.

20 At sub-human scales all other AT activities also operate without direct human judgement, so one execution algorithm can lead to many price-adjustment ticks in the same direction even without any feedback loops, leading to mini-flash crashes if the order book is not concurrently replenished. The vast majority of such events heal themselves through limit order book replenishment, leaving the crashes invisible at lower frequencies. If, on top of an execution algorithm at work, feedback loops form at sub-human scales, then the dynamics become nonlinear and the crash may well be visible at human scales, unless preemptive circuit breakers successfully kicked in.

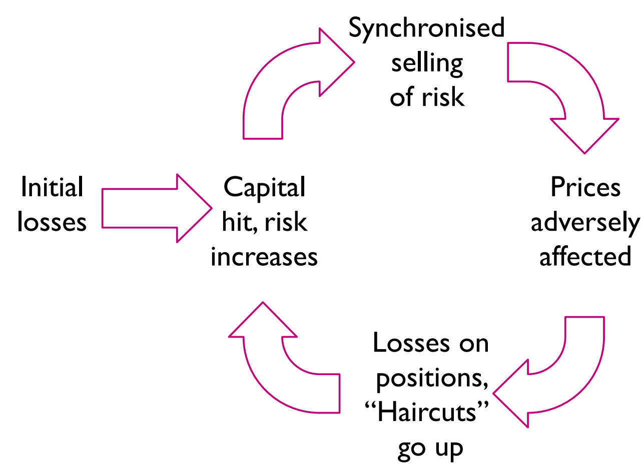
21 Also see DR29 (Annex D refers). Most of the evidence in that review is based on US data.

A better understanding is needed of the safety envelope within which such feedback loops either do not arise or can be interrupted with confidence and at low cost (see, for example, the case study of risk management in the nuclear energy industry in DR26[22]). Amongst the prudential devices mooted to reduce or cut feedback loops are various circuit breakers and trading limits that ideally limit false positives (breaks that prevent the market from adjusting to the new fundamental value following an informative piece of news) by being able to judge whether market stress is due to fundamentals or feedback loops. The role of circuit breakers is discussed in more depth in Chapter 6.

Risk feedback loop

Financial crises typically involve endogenous risk of the following sort (see Figure 4.8). Assume that some financial institutions are hit by a loss that forces them to lower the risk they hold on their books. 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 impact on 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[23]. Versions of this loop apply to HFT market makers: given the tight position and risk limits under which high frequency traders operate, losses and an increase in risk lead them to reduce their inventories, thereby depressing prices, creating further losses and risk, closing the loop. This value-destruction in turn can cause banks to stop performing their intermediation role with adverse spill over effects on the real economy. Devising measures to limit this loop *ex-ante* or to cut the loop *ex-post* is notoriously difficult, not least due to the fallacy of composition[24], although attempts to further diversify among market participants should be encouraged.

Figure 4.8: Risk feedback loop



22 DR26 (Annex D refers).

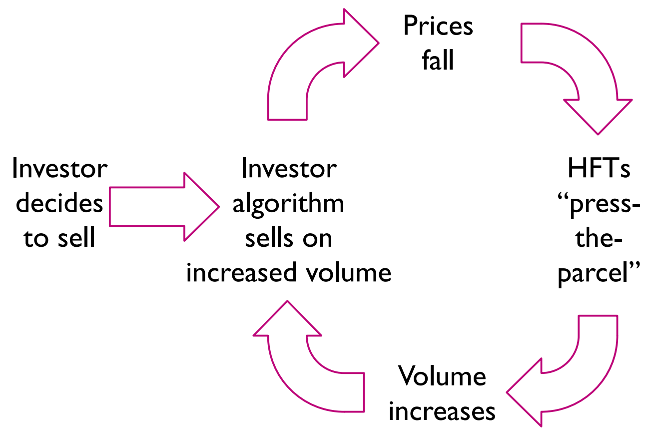
23 For further details see Brunnermeier & Pedersen (2009). That such feedback loops can lead to dynamics exhibiting fat left tails (i.e. an increased probability of extreme losses) as well as volatility clustering has been shown in both rational, forward-looking models (Danielsson *et al.* (2011)) as well as in models with myopic agents (see Danielsson *et al.* (2004), for a model where the friction generating the feedback loops is given by a value-at-risk constraint; and Thurner *et al.* (2012), where the friction is a margin constraint). For empirical evidence on the effect of losses or redemptions see Joshua & Stafford (2007). For evidence on how shocks can propagate across different financial institutions see Khandani & Lo (2007), and Khandani & Lo (2011).

24 This fallacy is discussed in Samuelson, P. (1947), and goes back at least to John Stuart Mill (1846). It refers to arguments that attribute a given property to the whole only because the individual elements have that property. In the present example, it is tempting to say that because each financial institution sheds risk and becomes safer, the financial sector as a whole must be safer. Quite the opposite is true.

Volume feedback loop

Whether the analysis in the official US Commodities and Futures Trading Commission/Securities and Exchange Commission (CFTC/SEC) report[25] of the Flash Crash events of 6 May 2010 turns out to be accurate and complete or not (see DR4 and DR7[26] for further discussion), it does illustrate a potential driver of risk. The CFTC/SEC report describes a possible scenario whereby some HFT algorithms may *directly* create feedback effects via their tendency to hold small positions for short periods. Such a 'hot-potato' or 'pass-the-parcel' dynamic occurred on 6 May 2010 when trading amongst high frequency traders generated very large volumes while hardly changing the overall net position at all (see Figure 4.9 where a sale leads to a price drop and an increase in HFT inventories which HFT then quickly try to reduce, leading to increased trading volumes, which in turn encourage the original algorithm to sell more). Because financial instruments were circulating rapidly within the system, the increase in volume triggered other algorithms which had been instructed to sell more aggressively in higher volume markets (presumably on the basis that higher volume means lower market impact), selling into the falling market and closing the loop[27]. Circuit breakers and algorithm inspection may prevent some of these loops developing.

Figure 4.9: Volume feedback loop



25 US Commodity Futures Trading Commission and the US Securities and Exchange Commission (2010).

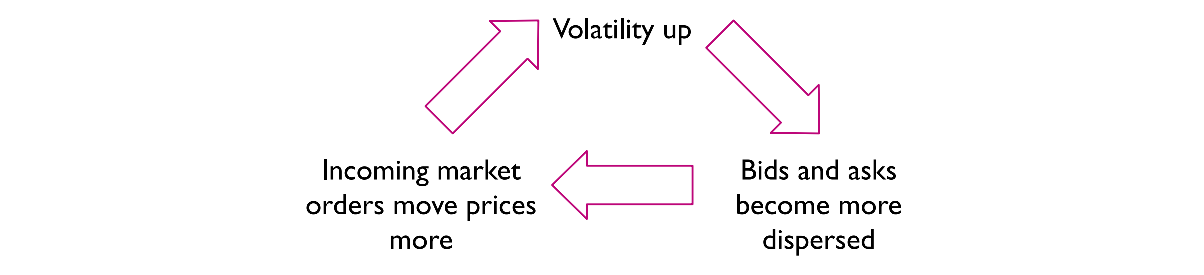
26 DR4, DR7 (Annex D refers).

27 Kirilenko *et al.* (2011) provides evidence that some market participants bought and sold from each other very rapidly 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 the need for a better understanding of certain CBT and HFT strategies and their interactions with each other and other market participants. While not able to directly identify CBT or HFT, Easley *et al.* provide evidence of the speed and magnitude of unusual trading behaviour in more securities during the Flash Crash, see Easley *et al.* (2011a).

Shallowness feedback loop

Closely related is the potential feedback loop described by Angel (1994)[28] and Zovko and Farmer (2002)[29] illustrated in Figure 4.10. Assume an initial increase in volatility, perhaps due to news. The distribution of bids and offers in the order book adjusts and becomes more dispersed[30]. 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). This increase in volatility in turn feeds back into yet more dispersed quotes, and the loop is closed.

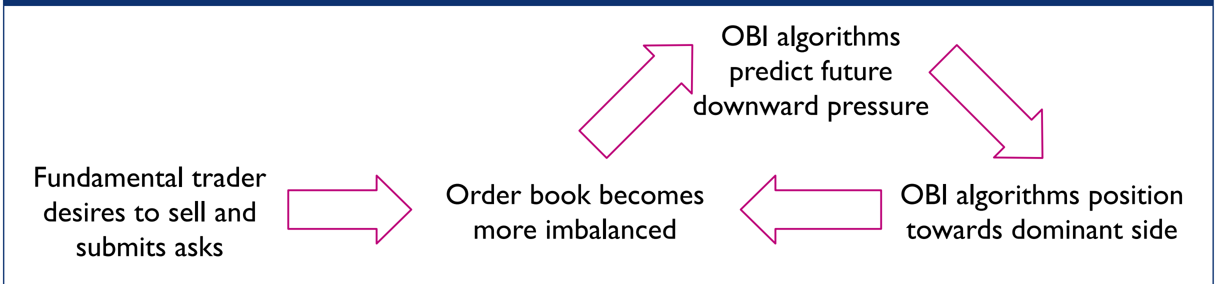
Figure 4.10: Shallowness feedback loop



Order book imbalance (OBI) feedback loop

Some HFT algorithms trade by relying on statistical tools to forecast prices based on the order book imbalance, roughly defined as the difference between the overall numbers of submitted bids and submitted asks. Assume that a trader wishes to lay off a large position and adds asks at the best ask. The order book now becomes imbalanced towards selling pressure, and high frequency traders operating an OBI strategy add their own offers and remove some of their own bids given their prediction of a falling price (the market falls without a trade, confirming the forecast), creating further OBI and closing the loop (see Figure 4.11). The feedback is interrupted once fundamental traders (i.e. traders with investment decisions based on fundamental valuation), who tend to be slower traders, step in to purchase, anchoring prices again to fundamentals.

Figure 4.11: Order book imbalance feedback loop



28 Angel (1994).

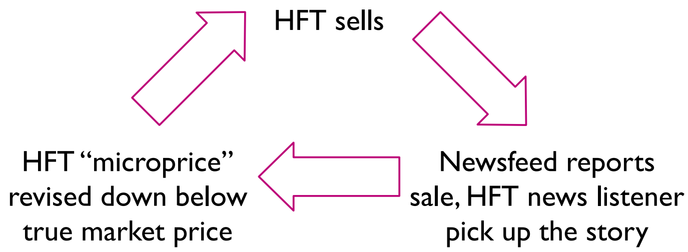
29 Zovko & Farmer (2002).

30 For an examination of these types of book changes prior to news about earnings, see Lee *et al.* (1993).

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 systems, is discussed in greater detail in Chapter 2 and in DR8[31]). High frequency traders buy or sell depending on where prices differ from their own perceived fair value; if the transactions of HFT systems are reported in news feeds, and picked up on by other HFT systems, those systems can be led to revise their price in a direction that encourages them (or other high frequency traders) to make similar trades (see Figure 4.12).

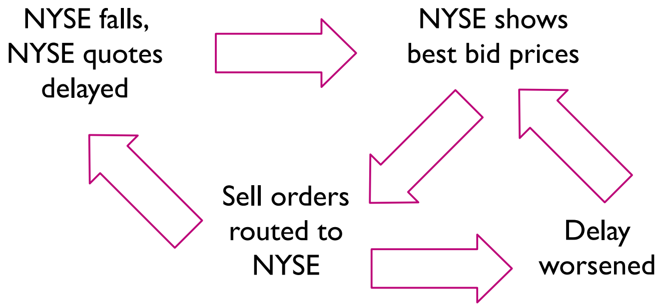
Figure 4.12: News feedback loop



Delay feedback loop

Eric Hunsader of Nanex Corp. suggests[32] the potential of the following very simplified feedback loop which may have operated during the Flash Crash. Consider a fragmented market suffering from overall selling pressure of a group of high-capitalisation stocks (for example originating from the sales of E-Mini futures), and assume that the NYSE quotes lag slightly (see Figure 4.13). 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. High frequency traders that follow algorithmic momentum strategies short those stocks and, given the unusual circumstances in the market, 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, so that the delay feedback loop amplifies the pricing feedback loop[33]. The provision of prompt public information on delays, possibly during a circuit breaker-induced pause, would reduce the occurrence and violence of such loops. Furthermore, some throttles, such as order-to-trade ratios, on excessive quote submissions that are intended to slow trading venue systems may prevent some of these delays in the first place.

Figure 4.13: Quote-delay feedback loop



31 DR8 (Annex D refers).

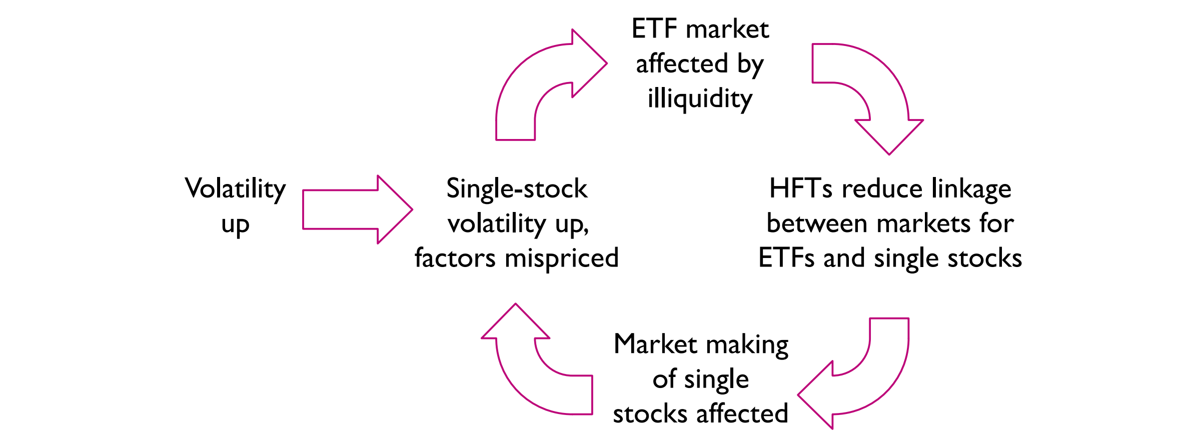
32 See DR7 (Annex D refers), and also: Hunsader (2010).

33 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 previous month.

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 makers to pause their market making activities. In response, the illiquid and stub ETF prices for indices provided false systematic factor signals upon which factor models price individual securities on a risk-adjusted fashion, feeding back into the pricing of individual securities, and thereby closing the loop (see Figure 4.14)[34].

Figure 4.14: Index feedback loop



While this chapter focuses on the sources of financial instability in modern CBT environments, CBT by itself can, of course, also be beneficial; for example, it can lead to stabilising dynamics. Consider the index feedback loop. When idiosyncratic volatilities create distorted valuations across venues and products in normal times, high frequency traders engage in profitable latency arbitrage by buying low and selling high. They thus align valuations across venues and products, which leads to welfare-improving allocations and prevents the dislocations of prices and the resulting loss of confidence in markets which could otherwise have become a fertile breeding ground for destabilising loops[35].

Notice that while the unconditional average volatility may be reduced by CBT due to these stabilising dynamics in normal times (and therefore much of the time), the existence of nonlinear reinforcing feedback loops that can create very large swings under exceptional circumstances prompts further caution about the simplistic use of volatility as a criterion for judging the health and safety of markets.

4.3.2 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 as 'normalisation of deviance' 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, on 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)[36] 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*.

34 Madhavan (2011) *op. cit.* analyses the feedback loop between ETFs and the underlying securities during the Flash Crash.

35 One piece of evidence that points towards the role of high frequency traders as removers of (transitory) pricing errors can be found in the working paper, Brogaard *et al.* (2012).

36 Perrow (1984).

As Cliff and Northrop note[37], the American sociologist Diane Vaughan has produced detailed analyses of the process that gave rise to the normal-failure losses of Challenger and Columbia space shuttles, and has argued that the key factor was the natural human tendency to engage in a process that she named[38] '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 a period of time if the price of that security moved by more than a given percentage within a short time. In response to the Flash Crash, the US SEC has now imposed 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 aware that they 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 reports (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 avoided on 1 September 2010, in a similar period when quote volumes exceeded even those seen at peak activity on 6 May 2010, but no official investigation was commissioned to analyse the events of September 2010. 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. The systemic effects of these failures may not have been fully considered.

Of course, the next flash crash will not be exactly the same as the previous one. 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 since the next market failure may well be a failure of risky technology that, like the Flash Crash, has no clear precedent.

Cliff and Northrop (DR4) argue that the normalisation of deviance poses a threat to stability in the technology-enabled global financial markets. Indeed, 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. In other words, it is 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 systems, with no single overall management or coordination. Such systems are so radically different 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

37 DR4 (Annex D refers).

38 Vaughan (1997).

Haldane (2011)[39]. A comparison between the systemic risk and safety in the nuclear industry and in finance can be found in DR26.

4.4 How can computer-based trading affect financial stability in future?

4.4.1 Forces at play

The strength of feedback loops depends on a number of variables, especially the capitalisation levels and the leverage ratios of financial institutions, and the degree of diversity of market participants. For instance, if liquidity is provided by lightly capitalised HFT operators rather than market makers with large inventories[40], then the passing-of-the-parcel scenarios as well as the resulting feedback loops are stronger because inventory management with little capital requires the rapid offloading of positions when prices fall, creating a volume feedback fall in prices. In that sense at least, substituting speed (with the cancellations that come with speed) for capital works well for market making purposes in normal conditions but may not work well in times of greater stress where the lack of capital can contribute extremely quickly towards positive instead of negative feedback loops[41, 42].

Similarly, the race for speed can reduce market diversity if market participants adopt the same (minimal number of) lines of code, leading to algorithmic crowding (see DR27). If similar mistakes creep into those algorithms because of inherent programming or engineering difficulties (DR26), the less diversified market participants may act in unison, creating stronger feedback loops, especially if the substitution of speed for capital further erodes investor confidence.

Moreover, diversity itself can worsen 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. These can lead to instantaneous synchronisation of actions among a group of institutions if they are all subject to the same regulations, constraints, and coordination mechanisms. For example, 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. Where liquidity provision is in the hands of a small number of large but lightly capitalised market participants, this can lead to a rapid drying-up of liquidity (see also DR7 for an estimate of the strength of this effect). In fact, informational asymmetries (where one participant knows that they know less than another) have the power to strengthen the endogenous risk mechanism in a variety of ways (for more details, see DR9[43]). This is because investors may mistake a temporary forced sale for a sale based on negative privileged information, leading to even lower prices.

Furthermore, information in CBT environments can present many subtle dangers. For example, even if all active market participants know that a certain event has not occurred, market prices and quantities would still not be able to completely discount the event, for even if it is known by all that the event has not occurred, there can still be participants who do not know that others know that it has not occurred. It can be argued that technology has further removed, to a certain extent, common knowledge from markets. This is also a recurring theme in the interviews of computer-based traders conducted by Beunza *et al.* (IN1)[44]. 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')[45]. This has, in effect, introduced further levels of complexity, if only because the outcome of the market now depends in a non-trivial

39 Haldane (2011).

40 For empirical evidence on the importance of liquidity providers' balance sheets see Comerton-Forde *et al.* (2010).

41 For further details on how intermediaries manage risk and how this affects price dynamics see Duffie (2010). For empirical measures of the magnitude see Hendershott & Menkveld (2011).

42 It seems that, in terms of market structure, the presence of low latency 'cream skimming' liquidity providers with little capital has driven out market makers with capital and inventories. Unfortunately, there is little empirical work available on how pervasive 'cream skimming' really is; see Menkveld (2012).

43 DR9 (Annex D refers).

44 IN1 (Annex D refers).

45 See: Rubinstein (1989), and Halpern & Moses (1990).

way on what each trader believes any other trader believes about all other traders' beliefs, and so on. The direct link between market outcomes and the fundamental events that ought to act as anchors for valuation has been further severed and replaced by a complex web of iterated and nested beliefs[46]. Meanwhile, there may have been a weakening of shared informal norms (such as the implicit sense of obligation to refrain from certain actions or to step in and help out). Thus, each actor has a very limited view of the entire landscape and limited interaction with other humans[47].

Interviews conducted by Beunza *et al.* (IN1) suggest that, as a result of these shifts, computer-based traders have learned to use social networks both in case of failure in the electronic networks and as a means to access information not available 'on the wire', relating to for example, as market sentiment or, more importantly, participation. According to Beunza *et al*. (IN1), they rely on such networks to form views on who is still trading, what are volumes, are people taking risks, are prop desks busy, or are they sitting on their hands waiting, trying to get a feel for markets[48]. The information gathered from such links in turn informs the trader's decision on whether to kill the algorithms.

This is not to say that full transparency is always *ex-ante* optimal. *Ex-ante* transparency can have significant costs in general, since risk-sharing opportunities are prevented, a phenomenon commonly referred to as the Hirshleifer effect[49]. In pre-trade, for example, transparency may increase market impact costs (also called implementation shortfall; for example, due to order anticipation or due to informational considerations) and therefore lead *ex-ante* to less demand for securities that are more difficult to trade, in turn making capital-raising by various entities and corporations more expensive, leading to less liquid markets, and so on. Given the current market structure, dark markets (markets without pre-trade transparency) have a role to play in normal times. In periods of stress, however, it may well be that the social costs arising from the absence of common knowledge are large enough to warrant an institutional mechanism designed to remove lack of transparency quickly in such periods and illuminate all markets, as suggested in DR9.

To summarise, computer-based markets may introduce an additional obfuscatory layer between events and decisions. This layer harbours the risk that unfounded beliefs or rumours may contribute to self-reinforcing cascades of similar reactions and to the endogenous risk, pushing valuations away from fundamentals.

4.4.2 Risks for the future

Left to its own devices, the extent of CBT may still grow, and the endogenous risk factors described above will continue to apply. Some aspects of CBT, however, may have started to find their level. For instance, there are natural bounds on the extent of trading that can be generated by proprietary ('short-termist') high frequency traders. First, those trades that have such traders on both sides of a transaction can generate profits only for one. Second, the no-trade theorem[50] would predict that trade will collapse once it becomes known that the only trades are those offered by short-termist proprietary traders who do not have an incentive to hold the securities for fundamental reasons. And third, much of the profits and rents of HFT are competed away under increasing competition. Recent reports suggest that profits of HFT companies have declined[51], and a study[52] in 2010 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 in London and EuroNext equity trading (see DR5). Practitioner views seem to confirm this

46 For a detailed analysis, see Shin (1996).

47 This partial view has the potential in turn to lead, in turn, to actions based on misinterpretations and to unintended consequences, since no actor can possibly have forecast, considered and internalised all the direct and indirect effects of their actions, including all reactions and reactions to reactions and so on, see DR11 (Annex D refers). The 'culture of restraint' amongst NYSE specialists that led members to follow the norms even if it was against their self-interest, as documented by Abolafia, (2001), would not be sustainable in a CBT environment where social norms become weaker and deviancy can set in, as in the 'broken window' theory. For example, in the case of quote-stuffing: if a floor trader had created deliberate noise by rescinding hand signals in previous decades he would probably have been rapidly stopped from doing so.

48 IN1, p. 18 (Annex D refers).

49 Hirshleifer (1971).

50 Milgrom & Stokey (1982).

51 See, for example, Cave (2010).

52 Kearns *et al.* (2010).

scenario[53]. Thus, CBT may gain market-share as more buy-side investors use it, but HFT, in the sense of proprietary intermediation trading, is naturally limited by the fundamental trading volume of real-money investors.

There are a few factors, however, that could further increase HFT. First, if proposals to remove some of the dark pool waivers are adopted, institutional investors may be forced to trade more on lit pools. Second, some public markets still retain an element of open-outcry, which may be eroded. Third, the progressive migration of some over-the-counter (OTC) markets towards clearing may ultimately lead to standardisation and trading on exchanges. Fourth, if market making activities by banks were to fall under proprietary trading, more of the market making would be done by unaffiliated and lightly capitalised high frequency traders. Finally, HFT will continue to expand into new public markets, including emerging markets and electronic retail market places such as today's Amazon or eBay[54]. To conclude, over the next five to ten years, there is likely to be further substitution of open-outcry, dark and OTC trading in favour of CBT, rendering all the effects described above more salient still.

One aspect apparent in the interviews by Beunza *et al.* (IN1) is the feeling of 'de-skilling' among market participants, especially on the execution side. As trading is done increasingly by algorithms, traders are being replaced by 'anti-traders' whose role it is to shut trading engines down as well as to tweak them, rather than to take decisions: "Traders gradually lose the instincts and tacit knowledge developed in floor-based trading"[55]. With fewer traders actively involved, and also with fewer observers of the markets due to automation, the chances of identifying risks early may be reduced.

Financial institutions optimise their trading 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 example, 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. This strongly suggests that if a testing facility for algorithms were to be introduced, the safety of individual algorithms is not a sufficient or indeed even a necessary criterion for systemic stability. It follows that, when thinking about the future of CBT and financial stability, assumptions have to be made as to the future regulatory and other constraints imposed upon markets and the new dynamics carefully assessed. For example, 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 which will affect future stability is the market segmentation between the various competing trading venues. Aligning valuations for single securities, multiple securities, and derivatives across venues is a socially useful task that high frequency traders 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 HFT will not be able or willing to perform this task in periods of market stress, margin calls or collateral pressure.

Non-linearities in liquidity provision (leading to quick reversals between feast and famine) are an important 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 an HFT world because of the effects discussed above, as well as the fact that algorithms are fine-tuned to the time-series on which they are tested and will be taken off-line when faced with sharp price movements, lack of liquidity or delays that render the models inapplicable. These considerations make the 'market makers' problem' of inventory and information management not only different but also altogether more difficult.

Finally, in a world with multiple trading and pricing venues that are interconnected by high frequency traders, the network topology determines the stability and the flow of information and trades. Together with company-owned dark pools (broker crossing networks), 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

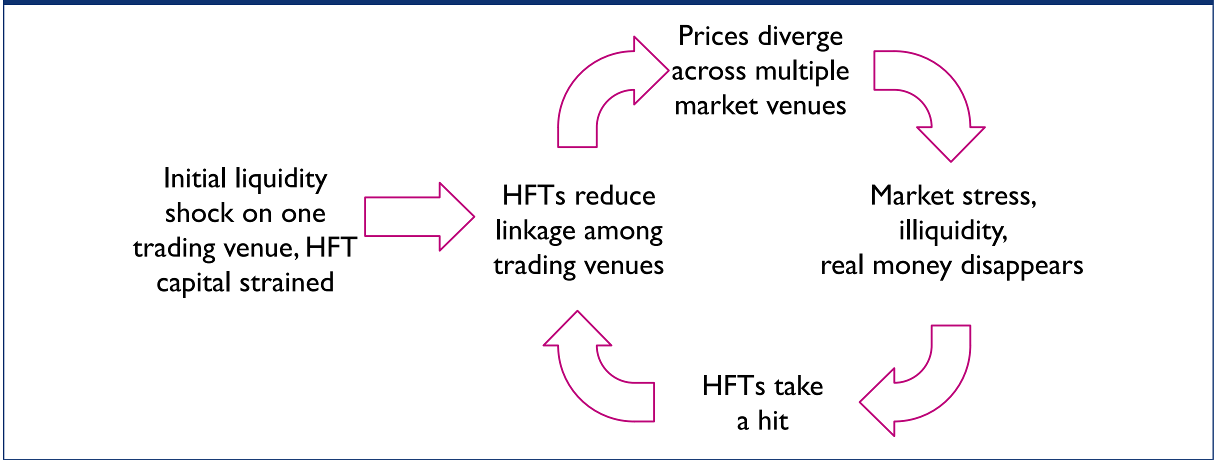
53 WR3 (Annex D refers).

54 See Jopson (2012).

55 IN1, p. 18 (Annex D refers).

HFT operators 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 high frequency traders buy low and sell high. But in stressed markets, their capital may be limited, or the traders themselves may start to doubt the prices (as happened during the Flash Crash) and refrain from arbitrage. Institutional investors then start to mistrust valuations across the board, and the resulting pressures mean that high frequency traders no longer contribute to liquidity provision, which makes price divergence across trading venues worse still. And so the shock is transmitted through the network, and its effects are reinforced by yet another positive feedback, as illustrated in Figure 4.15. Trades and transactions will occur at socially inefficient prices, and mark-to-market valuations can only be done to multiple and illiquid marks.

Figure 4.15: Systemic divergence feedback loop



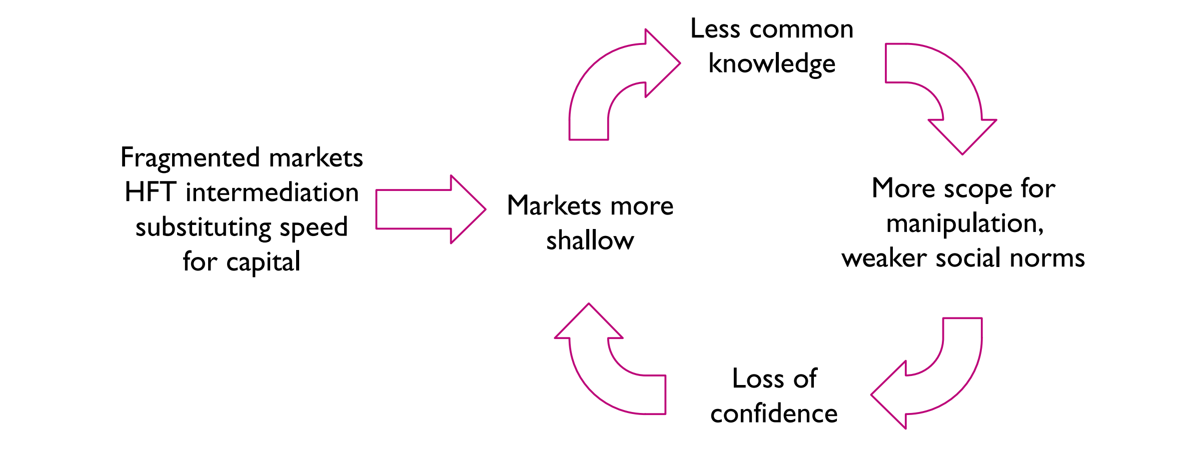
Understanding how to avoid such situations, and to contain them when they do occur, is a topic not only for further research but for urgent consideration by policy makers and regulators. As argued in Chapter 6, the reality and risks of multiple trading venues suggests that authorities, working together with industry and academics, should develop proposals for a system of cross-market circuit breakers. The degree of integration is open for debate. A loose collection of circuit breakers or limit up/limit down restrictions may in fact create greater endogenous risk. For example, when market prices on one security in one market are stuck on a limit down or limit up boundary while related prices on other trading venues are not, linkages are broken and price coherence made impossible, leading to the possible activation of the systemic divergence feedback loop. On the other hand, a break on a minor market triggering a break on a major venue can create unnecessary confusion and uncertainty.

Furthermore, interviewees in the study by Beunza *et al.* (IN1) stated that CBT actions during a dislocation to a primary market in a benchmarked price system will in part be determined by the outcomes of social communications during the event. Depersonalisation also may lead to weaker social norms and to less common knowledge that in turn can give rise to deviancy[56] and reduce investor (as well as market maker) confidence. The lack of confidence may lead to shallower markets, which in turn create less common knowledge and more opportunities for manipulation (see DR20[57] for further details on market abuse), see Figure 4.16.

56 For example, various presumed low latency market manipulation strategies that operate across markets, and which are therefore harder to pinpoint, have attracted attention. One possible such strategy (dubbed 'The Disruptor' by Nanex) would work as follows. It relies on the fact that arbitrage algorithms keep the E-Mini futures and the ETFs on the S&P500 in line. The strategy disrupts the E-Mini futures during a time where the ETF market is soft (or is made soft by the algorithm) by selling excessive E-Minis and expecting the other side to have been buying, and therefore it can forecast that the other side will hedge in the ETF market. The speed advantage of the originating algorithm means that it may purchase from the arbitrageurs at a profit-making price.

57 DR20 (Annex D refers).

Figure 4.16: Systemic market structure feedback loop



Shallower markets may increase the possibility of various feedback mechanisms and therefore systemic risk, and vice versa[58]. It would seem that policy-makers ought to take a step back from incremental regulatory adjustment and rethink the deeper fundamentals of healthy markets. For example, various proposed amendments to the market structure, such as Basel III and the bank ring-fencing proposals in the Vickers Report, may imply that broker-dealers are likely to reduce capital and funding liquidity allocated to market making, thereby leaving markets shallower and more vulnerable to endogenous risk[59].

4.5 Conclusions

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 beyond the capacity of human intervention, and in that sense an important speed-limit has been breached. Because of the speed advantages that computers can offer CBT 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)[60] than in the past; and the primary source of liquidity provision has changed, to computer-based and HFT systems, which has implications for the robustness of markets in times of stress.

Research thus far provides no direct evidence that HFT 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 as aeroplanes and spacecraft, can also affect the engineering of CBT systems.

The next chapter discusses market abuse in the context of CBT.

58 The evidence on the question of whether HFT leads to markets that are less deep is not fully conclusive. DR5 tentatively attributes to HFT part of the significant increases in FTSE 100 mean and median LSE order book depths displayed at the best quotes between January 2009 and April 2011, although the authors are silent on depth beyond the best quotes. They also show that FTSE small cap depths have improved much less. Using different realised impact measures, DR1 finds little change in realised impact over the same period. Using more recent data, Nanex finds huge variation in the depth of the first ten layers of the E-Mini futures contract with a possible downward path over the period from April 2011 to April 2012.

59 Shallowness operates directly through the shallowness feedback loop and the systemic market structure feedback loop, but in general it accelerates all loops that rely on price adjustments, such as the hedge feedback loop, the risk feedback loop, the volume feedback loop and so on.

60 The paper by Biais *et al.* (2011), focuses on asymmetric information arising not from the fact that some traders uncover new information but from the fact that their lower latency simply allows them to have foreknowledge about information that will be hitting markets a fraction of a second later anyway.

5 Market abuse and computer-based trading

Key findings

Economic research thus far provides no direct evidence that high frequency computer-based trading has increased market abuse. However, most of this research does not focus on the measurement of market abuse during the continuous phase of trading.

Claims of market manipulation using high frequency trading techniques are consistently reported by institutional investors (pension funds or mutual funds) internationally.

The same investors describe having little or no confidence in the ability of regulators to curb the behaviour they describe as abusive.

Policy makers should take such perceptions seriously, whether or not they are at variance with reality: the true extent of abuse will never be known and it is perceptions that determine trading behaviour and investment decisions.

Increasing the ability of regulators to detect abuse and producing statistical evidence on its extent would help to restore confidence.

5 Market abuse and computer-based trading

5.1 Introduction

The London Interbank Offered Rate (LIBOR) scandal supplies an excellent example of attempted price manipulation on a seemingly very large scale. It reminds us that forms of manipulation not conducted using electronic means are very much alive. However, a possible link between high frequency trading (HFT) and market abuse is frequently alluded to in the financial press, in regulatory reports and in direct interviews with investors. This link is the focus of this chapter: do increased latency differentials between agents increase the likelihood of market abuse and market manipulation in particular? Taking market abuse broadly to comprise insider trading and market manipulation, this chapter discusses several, not mutually exclusive ways in which HFT may matter to market abuse.

There is currently very little large-scale and data-based evidence on the incidence of abuse, whether in pre-HFT markets or today. However, qualitative evidence is accumulating that institutional investors perceive abuse as more widespread in today's markets, where HFTs may dominate liquidity supply. This perception has in itself the potential to affect the behaviour of slower and less-informed agents, particularly liquidity suppliers, with a consequent negative impact on market outcomes.

The concern that high frequency traders use a speed advantage over other agents to implement new abusive strategies generates much speculation, but other issues may be equally important. In particular, the growth of HFT may have altered the trading environment in ways that render some forms of abuse easier to perpetrate than in the past.

As a result, two courses of action seem available:

- Perception should be confirmed or corrected through the production of statistical evidence on the link between HFT and abuse.
- Significant investment should be made by regulators to acquire the skills and create the infrastructure that would allow them to deal with the surveillance issues that HFT has created and thus help to build the confidence of institutional investors.

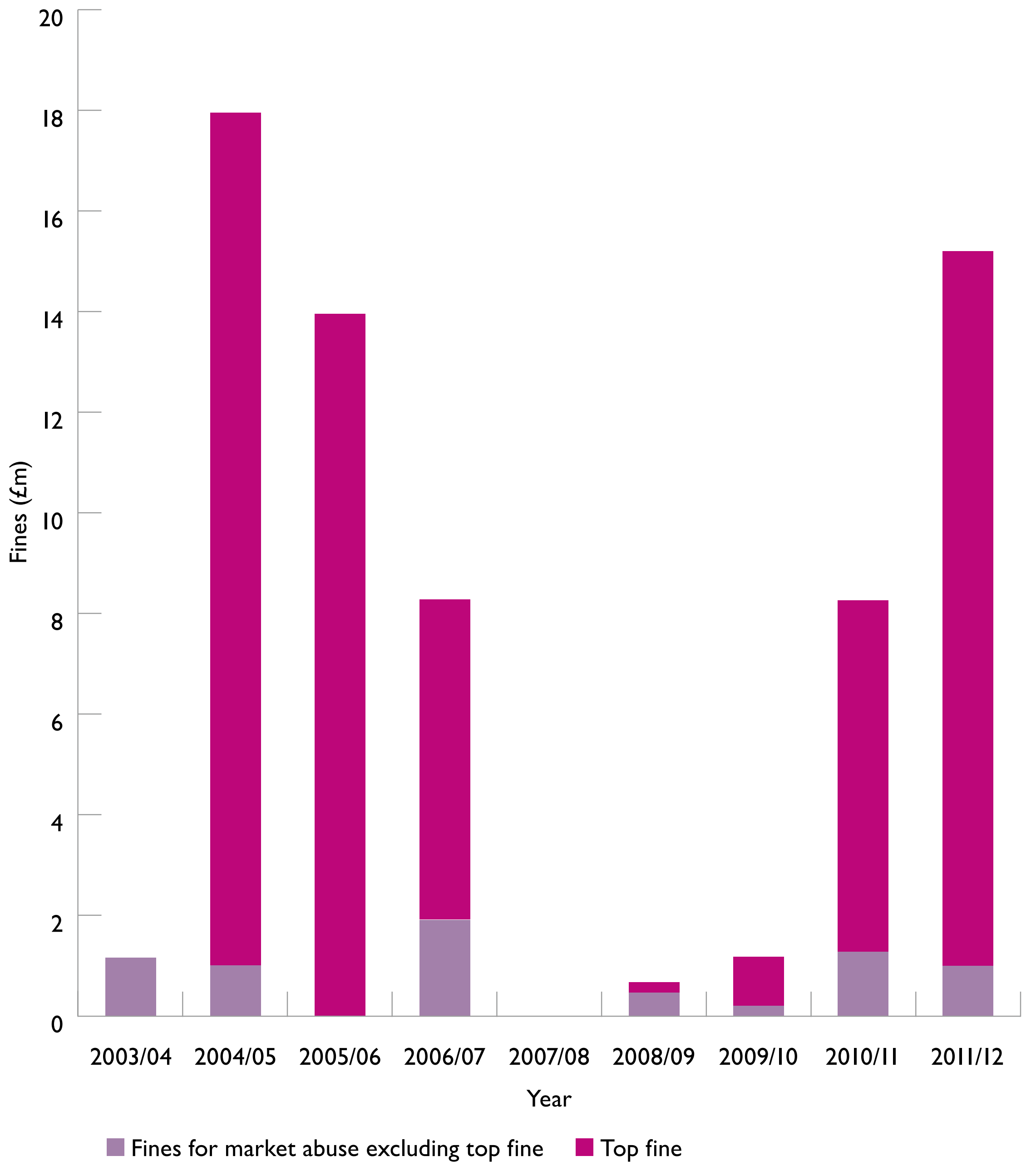
5.2 Past: what was the impact of computer-based trading on market abuse in recent years?

This section does not review large-scale and data-based evidence on the past incidence of abuse as such evidence simply is not available to academic researchers: abuse is rarely identified unambiguously and only a few cases every year are successfully prosecuted, even in the largest markets. For example, Figure 5.1 shows that the total value of fines for market abuse that have been levied on financial firms by the UK regulator since 2003 show year-to-year variability of the total fine and a general lack of trend. Where the total fine was large, it was dominated by a single massive fine[1]. No fines were levied in 2008[2].

1 Except in 2012 where three fines are clear outliers and make up the purple area of the bar.

2 It is important to note that because of the delays involved in detection and prosecution, the year in which the fine was imposed was typically not the year during which the related abuse was conducted.

Figure 5.1: Fines levied by the UK Financial Services Authority (FSA) for market abuse, 2003-2012



Source: FSA Annual Reports

None of the fines imposed by the UK Financial Standards Authority (FSA) so far appears to have been directly related to high frequency strategies, even though HFT has grown in importance over the past five years[3].

The USA has seen a very small number of cases of market abuse related to HFT. In late 2010, the US Financial Industry Regulatory Authority (FINRA), an industry body that has regulatory authority over its member firms, announced that it was fining Trillium Brokerage Services $1m, for order book layering. The relative lack of HFT-related cases is consistent with the interpretation that HFT is not giving rise to more abuse, or alternatively that such abuse is much harder to detect. It is certainly the case that the few penalties related to HFT pale into insignificance when compared to the fines imposed to date on the firms implicated in the low frequency LIBOR fixing scandal uncovered in 2012.

3 In May 2011, the FSA announced its intention to impose an £8m fine on Swift Trade, a now dissolved Canadian firm catering for day traders, for order book layering, but the outcome of this decision is still pending.

They included the largest single fines ever levied by the FSA and US Commodity Futures Trading Commission (CFTC). Barclays was fined £59.5m by the FSA, $200m by the CFTC and a further $160m by the US Department of Justice.

Other data also suggest that the fines for market abuse are relatively insignificant. Table 5.1 shows the number of criminal sanctions for market abuse imposed in some European Union (EU) Member States between 2006 and 2008[4]. The numbers are clearly low and indicate how unlikely successful criminal prosecution has been.

Table 5.1: Annual number of criminal sanctions for market abuse in various EU member states, 2006-2008

Member State	2006	2007	2008
Austria	N/A	N/A	21
Belgium	1	2	1
Cyprus	N/A	N/A	6
France	31	19	16
Germany	20	7	16
Italy	N/A	11	5
Luxemburg	0	0	0
Netherlands	1	2	4
Poland	4	11	8
Spain	N/A	14	11
UK	6	1	6

There is a widely held suspicion that the number of these cases is vastly underestimated. Although there is a lack of significant empirical evidence, regulators and economists have given much thought to the definition and implications of market abuse. These views are considered in the next section.

5.2.1 Regulatory and economic views of abuse

Manipulation can be achieved by various means, such as appearing informed (for example, bluffing), or through trading that temporarily pushes prices away from equilibrium, or by entering quotes that give a false impression of actual liquidity. Market abuse is distinct from securities fraud, which typically implies improper conduct on the part of a financial market professional towards their client. In cases of market abuse, abusers rarely know their counterparties and, even more importantly, abuse operates via the price process of securities.

In order to prosecute, regulators need to identify victims of abuse. The regulatory approach to abuse defines victims as those who have been penalised either through an unfair informational advantage or through artificially manipulated prices. Economists, on the other hand, study abuse because to them it has market-wide implications for liquidity, turnover, pricing efficiency and other variables affecting social welfare[5].

Both economists and regulators agree that the perception of abuse is crucial. One obvious reason for this is that the true incidence of abuse can never be established, and current and potential investors must therefore form an estimate of the likelihood of losing out unfairly against other counterparties. As a result, a high perceived likelihood of suffering a loss through interaction with other traders may have damaging implications for liquidity, efficiency and transparency. Abuse has obvious direct costs for its immediate victims, who suffer reduced investment performance via increased trading costs. But perceptions of abuse also affect the behaviour of liquidity suppliers who will protect themselves against

4 European Commission Staff Working Paper (2011a).

5 This is why economists are not concerned with securities fraud, which they consider a purely regulatory and legal matter.

expected trading losses by quoting poorer prices – the equivalent of an insurance premium. As a result, increased trading costs and poorer performance of investment and pension funds may spread to most other participants.

Another concern is that perceived abuse can reduce the equity market participation of large and less-informed investors who have a longer investment horizon. Perceived abuse may also have an impact on the market structure by influencing investors' choice of modes of trade. It may give incentives to large and long-term investors to shun centralised, transparent markets favour of 'private' execution of their trades, harming market transparency and price formation[6].

An illustration of this appears in the International Organisation of Securities Commissions (IOSCO) (2011):

> Some market participants commented to regulators that a loss of confidence in the fairness of markets may result in them reducing their participation in lit markets, or in shifting their trading into dark pools. If this effect is indeed occurring on a significant scale, it may result in a growing concentration of trading (at least in the lit markets) amongst a relatively limited set of market participants[7].

Clearly then, perceptions of abuse should not be dismissed.

5.3 Present: what is the current impact of computer-based trading on market abuse?

The commissioned review DR28[8] documents the changes in order flow and market structure that have taken place in the past decade and points to the possible risks that this may raise for market abuse and the potential for perpetrators to hide their actions. Preliminary empirical evidence is presented in three reviews commissioned for this report (DR22, DR24 and DR28)[9]. These studies use publicly available data to construct proxies for HFT activity and abuse.

The effects of HFT on end-of-day price manipulation on a large sample of exchanges around the world are tested directly in DR22 and the authors conclude that the data demonstrate that HFT leads to lower incidence of manipulation. Another study, using data from the London Stock Exchange (LSE) and Euronext Paris from the period 2006–2011 also reports a significant negative relationship between HFT activity and end-of-day price manipulation (DR24). A longer span of Euronext Paris and LSE data are used in DR28 to verify that HFT improves market efficiency and integrity. Overall, these studies provide a useful starting point from which to argue that fears of HFT generating abuse are unfounded. However, the studies focus, in the main, on a very specific part of the trading day (the market close) and do not attempt to measure the extent of market abuse during the continuous phase of trading.

Other forms of evidence have recently begun to appear. This evidence can be described as qualitative or 'soft' in that it does not allow the quantification of the reality of abuse as such but pertains entirely to perceptions of it.

6 In this respect, institutional investors must not be pitted against retail investors. The former often represent the pension and investment savings of households, and any trading costs suffered by large funds will ultimately be borne by private savers.

7 International Organization of Securities Commissions (2011), p. 29.

8 DR 28 (Annex D refers).

9 DR22, DR24 and DR28 (Annex D refers).

5.3.1 Evidence on the perceptions of abuse

Evidence on investor perceptions of abuse has been published in the specialised press. It can also be extracted from reactions to regulatory consultations and calls for evidence. Even more useful, because they are conducted on a larger scale, are several recently conducted surveys of institutional investors, which either touch upon or explore the link between HFT and abuse[10]. This chapter draws on several surveys in which large investors claim that the liquidity supplied by HFT is thin, unbalanced and often even 'illusory'[11]. Many respondents were critical of HFT, some vehemently so. For example from the Kay Review: "The majority (…) doubted that the liquidity which high frequency trading claimed to provide was real"[12].

From this, a claim often follows that such liquidity is not just the result of the nature of HFT-dominated order flow but in fact amounts to deliberate manipulation:

> There is broad agreement among traditional investors that abusive trading practices need to be tackled, with most concerns centring on the high number of computer-generated orders being used to manipulate 'real' liquidity[13].

A survey of large investors and HFT conducted for the Swedish Government similarly notes that:

> concern for market abuse is considerable (…) The majority of companies that were surveyed expressed concern that a large portion of high frequency trading was being used to manipulate the market. There are clear apprehensions that market abuse has become more extensive and difficult to identify as a result of the sharp increase in the number of orders and trades[14].

Similar conclusions were drawn in analysis of the responses to the consultation of the IOSCO, which contains two questions directly relevant to abuse and HFT[15]:

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

All the entities that responded "no" to Q9, indicating that they felt inadequately protected from new forms of abuse inflicted by HFT, were bodies representing the buy side. Similarly, the institutions representing the buy side answered overwhelmingly "yes" to Q10 (followed by description of specific concerns).

Therefore, while suggestive, the evidence appears strikingly consistent in its flavour, even though it originates from different sources and regulatory areas: large investors believe that there is a clear link between abuse and high-speed trading techniques. It can almost be read as a call for action on the part of regulators: for example from the survey SRI commissioned by the Project "Traditional investors agree that creating a level playing field and increasing efforts to prevent market manipulation should be prioritised"[16].

10 These investors, for example pension and investment funds, are collectively known as the 'buy side', following US terminology. This reflects the fact that they 'buy' financial assets with the savings that households entrust to them.

11 The surveys this chapter draws from include: Kay (2012); Finansinspektionen (Swedish Government) (2012); SR1 (Annex D refers).

12 Kay (2012), p. 29.

13 See SR1, p. 6 (Annex D refers).

14 Finansinspektionen (2012), p. 3.

15 IOSCO (2011).

16 SR1, p. 86 (Annex D refers).

More recently still, exchanges and high frequency trading firms have started to acknowledge the issue publicly. The Chief Executive Officer of NYSE Euronext stated that:

> The public has never been more disconnected, (…) the citizenry has lost trust and confidence in the underlying mechanism. (…) What used to be an investors' market is now thought of as a traders' market (…) We convinced ourselves (…) that speed is synonymous with quality and in some cases it might be. In other cases, it clearly isn't[17].

It should be noted, however, that more specific interpretation of such qualitative evidence is difficult, as there seems to be no clear agreement on the specific forms of abuse that might occur. The abusive strategies evoked include 'quote stuffing', 'order book layering', 'momentum ignition', 'spoofing', and 'predatory trading' (sometimes referred to as 'liquidity detection' or 'order anticipation')[18]. SR1 observes:

> Respondents largely agree that it is difficult to find evidence for market abuse or quantify its impact on their trades[19].

There is also a lack of agreement on what regulators should actually do. From the Kay report again:

> While the tone of submissions was hostile to HFT, there were few suggestions as to how the volume of such trading might be reduced or the activities of high frequency traders restricted[20].

And again from SR1:

> Views diverged considerably across the investor community as to which policy options are appropriate[21].

End investors are aware that there is a trade-off between curbing HFT abusive practices and discouraging legitimate computerised liquidity supply. Therefore the evidence on the perception of abuse seems compelling in its consistency across sources, perhaps a sign that something potentially damaging could be at work. But it is also less than helpful to policy makers because agreement is much weaker when it comes to defining precisely what is wrong or what measures regulators ought to put in place.

5.3.2 Interpretation: three scenarios

The possibility for elevated abuse arising from HFT activities is consistent with three scenarios, where the reality or the perception of abuse can be affected. Significantly, these scenarios are not mutually exclusive.

In Scenario I, high frequency traders take advantage of slower agents through a speed advantage. This may involve several of the strategies mentioned above. This corresponds to the claims consistently emanating from the buy side. It is worth noting that:

- most of these strategies are not new and do not require low latency as such (the obvious exception being quote stuffing, whereby a firm floods electronic systems with order messages to swamp other participants' systems with noise), but they may be more profitable using speed;
- for the most part, these strategies would be unambiguously classified as manipulative whether under the EU's Market Abuse Directive or its successors the Directive on Criminal Sanctions and the Market Abuse Regulation, since they quite clearly create 'false markets', 'false impressions of the demand for and supply of securities', and so on. Legislators, least in current draftings of the new European directives, have indeed taken the view that there is no obvious need to adapt existing definitions of abuse to an HFT context.

17 Sloan (2012).

18 Explanations of market abuse strategies can be found in Annex C.

19 SR1, p. 71 (Annex D refers).

20 Kay report (2012), p. 29 *op. cit.*

21 SR1, p. 25 (Annex D refers).

Under Scenario II, the growth of HFT has changed the trading environment, in particular the nature of the order flow, in ways that facilitate market abuse or increase the perceived extent of abuse. This covers situations which are different to the confrontation of 'fast versus slow agent' described above, where, for example, high frequency traders prey on each other and slower traders may also find it easier to conduct abuse.

The widespread claims of 'predatory trading' heard from institutional investors could be a prime example of this scenario[22]. Computer-based order flow may generate shallower and more predictable order flow that may be easier for a 'predator' (slow or fast) to manipulate because price pressure effects are easier to exert. This behaviour, especially if conducted through liquidity supply by temporarily modifying quotes following detection of sustained attempts to buy or sell the asset, could also be considered manipulative, as it attempts to profit from artificially pushing prices away from the levels warranted by their 'true' demand and supply. This statement illustrates:

> You can press the button to buy Vodafone, say, and have it executed in a second but in that period 75% of the liquidity has disappeared and the price has moved[23].

Scenario II may comprise other situations that can be generated by HFT, which have implications for abuse but do not in themselves constitute abuse. Rather, they are effects brought about by high frequency traders by dint of sheer speed. Among those is the HFT-related data 'overload' that almost certainly makes market surveillance vastly more difficult. Other examples are given in DR20.

Finally, Scenario III is one in which other market developments that have accompanied the growth in HFT (but only partly or not at all brought about by it), may also have contributed to an increase in the perceived or actual prevalence of abuse. These confounding factors are part of the new trading environment and may be hard to disentangle from HFT empirically, even though they are conceptually very distinct.

A prime example of this may be the fragmentation of order flow. Other examples include reduced pre-trade transparency tied to an increased use of 'iceberg orders' and dark venues, and the exceptionally uncertain economic environment since late 2007 which has led to high volatility and low liquidity. From SR1:

> The macroeconomic environment is seen as the key instigator in the decline in liquidity[24].

These developments contribute to a sense of unfairness and alienation from the trading process and, in some cases, may also facilitate abuse. Other illustrations are given in DR20.

The first Scenario is the one that generates the most speculation but Scenarios II and III may be equally important.

Can HFT make abuse *less* likely? This view rarely seems to be entertained, but high frequency arbitrageurs have powerful means and incentives to keep prices in equilibrium and to inflict losses on manipulators, as they focus on the relative pricing of economically similar securities.

5.4 Future: how might the relationship between computer-based trading and market abuse evolve in the next ten years?

Concerns have been consistently expressed about the extent of abuse perceived to be conducted through computer-based techniques, and these perceptions may have damaging consequences for market functioning. Because we do not know the true extent of abuse, they should not be dismissed as divorced from reality. What are the available courses of action?

22 Similar or identical strategies are called "liquidity detection" and sometimes "front-running" although the latter expression is inadequate in this context.

23 Paul Squires, head of trading for AXA Investment Managers, (The Economist) 25 February 2012.

24 SR1, p. 51 (Annex D refers).

5.4.1 Leave it to market mechanisms

It is the case that 'defence mechanisms' are being developed by the private sector to address perceptions of market abuse by high frequency traders. In the last quarter only, ingenious new trading systems have been launched specifically for institutional investors worried about predation, gaming and other strategies. These systems may feature one form or another of reduced pre-trade transparency, or they may reintroduce non-anonymity in trading, thus allowing a large investor to select the set of counterparties they wish to interact with, at the exclusion of, for example, proprietary traders. Other 'solutions' include software that assesses the 'toxicity' of order flow on each venue by examining patterns in recent activity and giving a 'green light' when trading is considered safe[25]. It has also been argued that dealing with HFT-driven abuse could be left to trading venue operators themselves, as they should have strong commercial incentives to provide order books that are seen as 'clean'.

However, such a purely market-driven approach may not adequately address perceptions of abuse for at least two reasons. First, many large investors doubt that trading venues have adequate incentives to be tough with high frequency traders, as HFT firms are one of their most important customer groups. Leaving it to trading venue operators to police abuse thus runs the risk of harming price formation by pushing investors to trade 'over-the-counter' (OTC) or in dark venues. This in turn increases the fragmentation of order flow across venues and aggravates order flow imbalances and volatility. Second, abuse is now possibly easier to conduct across venues and borders than in the past. Thus, market surveillance that puts too much reliance on individual trading venues is very unlikely to be effective. As a trader notes:

> we monitor the use of our infrastructure to prevent market abuse. But a high-frequency, multi-broker, multi-exchange strategy could trade buy orders with us, for example, and the sell orders elsewhere. If we don't see the sell order, it is very difficult for us to identify that instance of market abuse[26].

5.4.2 Address perceptions of the incidence of abuse

This strategy may cover two courses of action that are under active consideration in other mature markets and ought not, in principle, to be controversial. The details of implementation and the investment required are not trivial, however.

The first course of action is to convince market participants that regulators have the ability to deter abusers. This has two aspects: regulators must be able to levy penalties that are financially significant and they must also convince participants that the probability of abuse being detected and fines being incurred is not negligible[27]. Regarding the first aspect, the imposition of large fines in a few high-profile cases does not seem to have convinced investors that abusers have been deterred. Regulators seem determined to move on to the next level: making abuse a crime. This is most obviously the case in Europe under the proposed Directive on Criminal Sanctions for Insider Dealing and Market Manipulation. As for the second aspect, the market must also be convinced of the regulators' ability to detect abuse. The qualitative evidence available suggests that there is considerable room for improvement, as large investors express a lack of trust in the ability of regulators to tackle abuse. From SR1:

> About 90% of respondents do not believe that regulators have sufficient data, technology or expertise to effectively detect market abuse[28].

---

25 Order flow on a venue may be described as 'toxic' when, if a market participant executes on that venue, the probability of incurring a large trading loss is high. The participant might incur such losses when trading counterparties have superior information or are able to anticipate the individual's trading activity.

26 Holley (2012).

27 For a summary of the economics of corporate crime, including market manipulation, see The Economist (2012).

28 SR1, p. 24 (Annex D refers).

This in turn may comprise at least two areas for improvement:

- The generation and storage of data that would permit both real-time surveillance and *ex-post* investigations, and allow identification of firms and, in some cases, clients. This is being discussed at European level in continuing debates on transaction reporting, and in the USA, under the perhaps more ambitious header of 'Consolidated Audit Trail' (see EIA17 for more details)[29]. Requiring HFT firms to keep their trading code for some period, including a description of the rationale behind it, could also enable better enforcement.
- Once the data are available, there is a need to increase the regulators' ability to process and interpret them. In the USA, a report by the Boston Consulting Group recently warned that the Security and Exchange Commission (SEC) had to substantially improve its sophistication in information technology to stand any chance of detecting market abuse[30]. The report notes that:

> The agency does not have sufficient in-house expertise to thoroughly investigate the inner workings of such [trading] algorithms (…) [and recommends that] The SEC should have staff who understand how to (…) perform the analytics required to support investigations and assessments related to high-frequency trading. (…) the SEC requires people who know how high-frequency traders use technology[31].

These issues affect EU regulators to the same extent. Individual brokers can contribute to market surveillance (by, for example, filling out Suspicious Transaction Reports (STR) introduced in the EU by the Market Abuse Directive of 2003). One level up, so can exchanges and trading venues, which operate a mix of real and delayed time abuse/manipulation surveillance systems. As discussed above, exchange-level surveillance will be particularly useful with respect to strategies that are focused on a single order book, for example quote stuffing or order book layering[32]. But it seems generally accepted that the hardest part of the task falls to regulators, who must implement monitoring and surveillance across classes of assets, across exchanges and across borders. This approach to curbing abusive practices entails both international coordination and substantial investment but these efforts should both reassure large investors and deter potential abusers.

The second course of action to affect perceptions of abuse is to correct or confirm them through the production of statistically significant empirical evidence. HFT activity leaves an electronic trail so, given the appropriate data, researchers could generate at least preliminary evidence on the prevalence of simple patterns consistent with abuse. Such evidence could also inform policy making, as the three scenarios discussed earlier may have very different regulatory implications. At one end of the spectrum, if abuse is chiefly conducted by high frequency traders taking advantage of slower agents, then ways of slowing down markets may be considered. At the other end, if abuse is largely a matter of perceptions of large investors who feel alienated from market processes as a result of a new trading environment made up of speed, fragmentation and/or reduced transparency, then perceptions should be changed.

More large-scale statistical evidence will both help with perceptions and guide regulatory action. As SR1 concluded:

> any regulation of AT or HFT should not be undertaken without substantially greater analysis and a better understanding of the issues by policy makers and regulators. Some believe that extensive research is required to determine whether there actually is a problem in the market stemming from AT or HFT[33].

29 Very difficult technical issues are involved. An example is to synchronise clocks to the microsecond across the systems of different electronic venues to ensure that orders and trades can be reliably sequenced, a pre-requisite for any investigation of trading patterns. See EIA17 (Annex D refers).

30 Boston Consulting Group report (2011).

31 *Ibid.* p. 53; p. 260.

32 The London Stock Exchange disciplined one of its members for order book layering in 2009. See LSE Stock Exchange Notice N33/09.

33 SR1, p. 27 (Annex D refers).

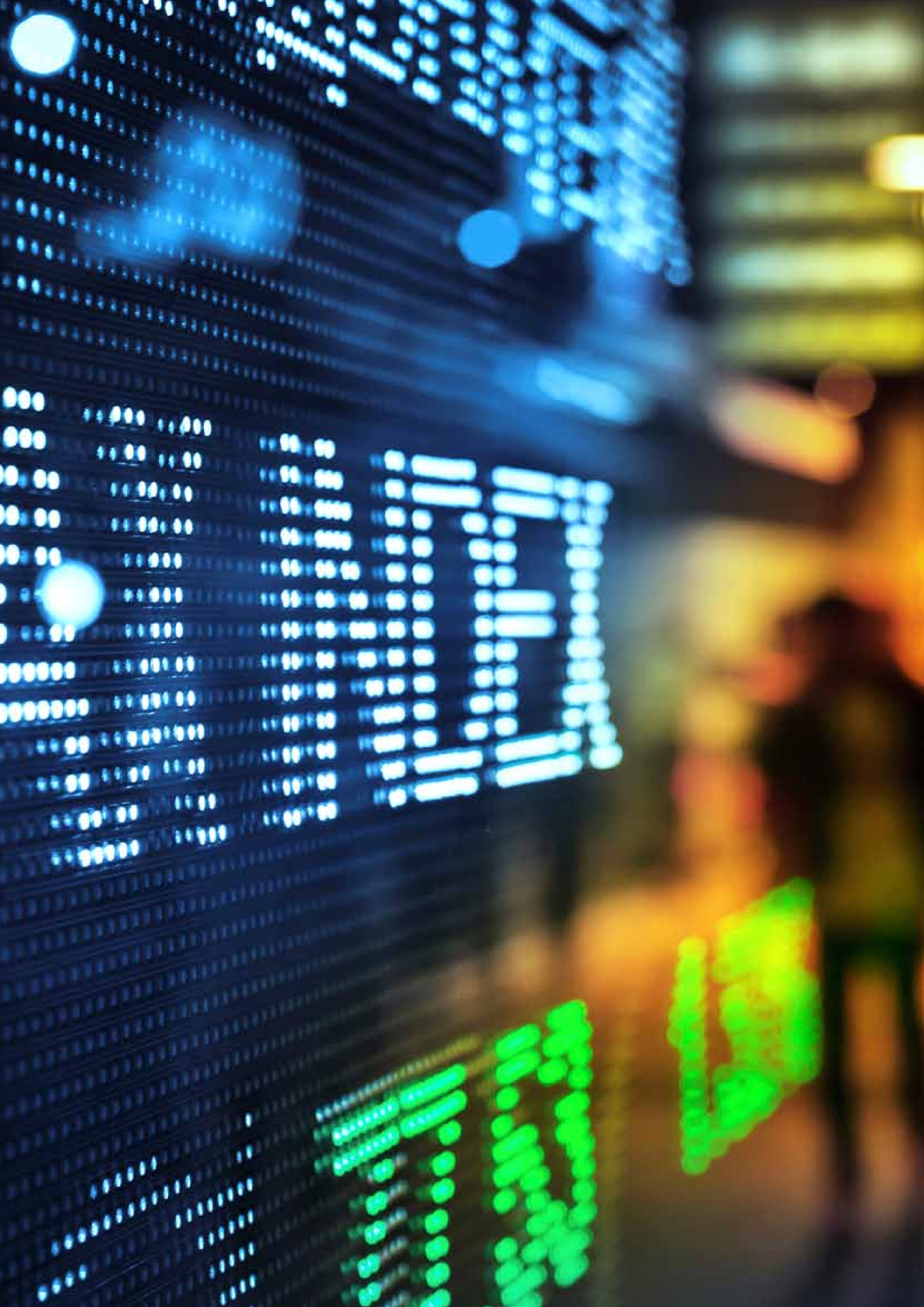
5.5 Conclusions

Economic research thus far provides no direct evidence that high frequency computer-based trading (CBT) has increased market abuse, although this research is at an early stage and incomplete: its main focus is not on the measurement of market abuse during the continuous phase of trading. However, claims of market manipulation using HFT techniques are consistently reported by institutional investors (pension funds or mutual funds) internationally. While these claims are not substantiated by evidence, plausible scenarios can be constructed which show how such abuse could potentially occur.

Policy makers should take such perceptions seriously, whether or not they are true: the actual extent of abuse can never be known and it is perceptions that determine trading behaviour and investment decisions.

Currently, surveys show that institutional investors have little or no confidence in the ability of regulators to curb the behaviour that those investors describe as abusive. While market mechanisms will provide some help, it is likely that these concerns will need to be addressed more directly. This chapter has argued that the need for direct intervention in market operation is not demonstrated at this stage, but that increasing the ability of regulators to detect abuse and produce statistically significant empirical evidence on its extent would help either to confirm concerns or lay them to rest, and thus restore market confidence.

The Report has considered a number of aspects related to CBT and outlined some of the challenges that CBT presents. In the next chapter, the report considers policy measures that aim to address those challenges and what economic impact those measures might have.



6 Economic impact assessments on policy measures

Key findings

Computer trading has changed markets in fundamental ways, not the least of which is the speed at which trading now occurs. There are a variety of policies proposed to address this new world of trading with the goals of improving market performance and reducing the risks of market failure. These policies include notification of algorithms, circuit breakers, minimum tick size requirements, market maker obligations, minimum resting times, minimum order-to-execution ratios, periodic auctions, maker-taker pricing, order priority rules and internalisation. The Foresight Project has commissioned a variety of studies to evaluate these policies, with a particular focus on their economic costs and benefits, implementation issues and empirical evidence on effectiveness. This chapter summarises those findings.

The key findings relating to the different policies are as follows, starting with those which were most strongly supported by the evidence:

- Overall, there is general support from the evidence for the use of circuit breakers, particularly for those designed to limit periodic illiquidity induced by temporary imbalances in limit order books. Different markets may find different circuit breaker policies optimal, but in times of overall market stress there is a need for coordination of circuit breakers across markets.
- There is also support for a coherent tick size policy across similar markets. Given the diversity of trading markets in Europe, a uniform policy is unlikely to be optimal, but a coordinated policy across competing venues may limit excessive competition and incentivise limit order provision.
- The evidence offers less support for policies imposing market maker obligations. For less actively traded stocks, designated market makers have proven beneficial, albeit often expensive. For other securities, however, market maker obligations run into complications arising from the nature of high frequency market making across markets, which differs from traditional market making within markets. Many high frequency strategies post bids and offers across correlated contracts. A requirement to post a continuous bid-offer spread is not consistent with this strategy and, if binding, could force high frequency traders out of the business of liquidity provision. Voluntary programmes whereby liquidity supply is incentivised by the exchanges and/or the issuers can improve market quality.
- Similarly, minimum resting times, while conceptually attractive, can impinge upon hedging strategies which operate by placing orders across markets and expose liquidity providers to increased 'pick-off risk' if they are unable to cancel stale orders.
- The effectiveness of the proposed measure to require notification of algorithms is also not supported by the evidence. The proposed notification policy is too vague, and its implementation, even if feasible, would require excessive costs for both firms and regulators. It is also doubtful that it would substantially reduce the risk of market instability due to errant algorithmic behaviour, although it may help regulators understand the way the trading strategy should work.
- An order-to-execution ratio is a blunt policy instrument to reduce excessive message traffic and cancellation rates. While it could potentially reduce undesirable manipulative trading strategies, beneficial strategies may also be curtailed. There is insufficient evidence to ascertain these effects, and so caution is warranted. Explicit fees charged by exchanges on excessive messaging and greater regulatory surveillance geared to detect manipulative trading practices may be more effective approaches to deal with these problems.

Key findings (cont.)

- The issue of maker-taker pricing is complex and is related to other issues like order routing, priority rules and best execution. Regulatory focus on these related areas seems a more promising way of constraining any negative effects of maker-taking pricing than direct involvement in what is generally viewed as an exchange's business decision.
- The central limit order book is of fundamental importance and everyone involved in equity trading has an interest in changes that improve the performance of the virtual central limit order book currently operating in Europe. The debate on how best to do so continues, in part because the opposing trends towards consolidation of trading venues and fragmentation is likely to remain unresolved in the near future.
- Internalisation of agency order flow, in principle, benefits all parties involved, especially where large orders are involved. However, the trend away from pre-trade transparency cannot be continued indefinitely without detrimental effects on the public limit order book and price discovery.
- Call auctions are an alternative trading mechanism that would eliminate most of the advantage for speed currently present in electronic markets. They are widely used already in equity markets at open and close and following a trading halt. But no major market uses call auctions exclusively to trade securities. To impose them as the only trading mechanism seems unrealistic as there are serious coordination issues related to hedging strategies that make this undesirable.

6 Economic impact assessments on policy measures

6.1 Notification of algorithms

6.1.1 The measure and its purpose

Algorithmic trading (AT) involves the use of computer programs to send orders to trading venues. Such algorithms now have widespread use among all classes of investors[1], and AT comprises the bulk of trading in equity, futures and options markets. AT is also fundamental to high frequency trading (HFT) strategies. A concern with AT is that an errant algorithm could send thousands of orders in milliseconds to a market (or markets), resulting in major market upheaval. Markets in Financial Instruments Directive (MiFID) II Article 17(2) proposes that investment firms engaging in AT must provide annually to the regulator a description of their AT strategies, details of the trading parameters and limits, the key compliance and risk controls which are in place, and details of how its systems are tested. The purpose of this measure is to ensure that algorithms used in trading are subject to proper risk controls and oversight.

6.1.2 Benefits

If descriptions were able to prevent unsound algorithms from operating in live markets, then this would be a measure contributing to the maintenance of orderly markets. Requiring firms to have demonstrated risk controls in place might make aberrant algorithms less likely to occur. If regulators require increased testing at the firm level for algorithms it suspects are flawed, fewer events affecting market liquidity due to malfunctioning algorithms might occur.

One additional benefit is that regulators will have to acquire greater technical sophistication to understand and evaluate the algorithms being used in trading, which would improve their ability to investigate abusive practices. However, this would require substantial increases in personnel and greater investments in technology.

6.1.3 Costs and risks

There are substantial costs connected with meeting notification requirements in general, and particularly as currently stated in Article 17(2). Cliff (EIA16)[2, 3] argues that just providing a full description of an AT strategy requires not only all the programs that have been written to implement it, but also the full details of the code libraries used, as well as the software tools involved. Moreover, these descriptions must include the actual computations required, the algorithms that affect the computations, and full details of how the algorithms are implemented. Providing information on the other aspects of the proposed regulation would be similarly complex. Regulators, in turn, would then have to analyse this material, and determine what risk, if any, this algorithm poses to the market. This would require substantial expertise at the regulator level with complex computing systems and analysis. An additional risk to consider is that algorithms are updated frequently, meaning that annual reviews will be ineffective in actually capturing the risk facing the markets.

1 The survey SR1 (Annex D refers) commissioned by the Foresight Project found that algorithmic trading is used by 95% of asset managers, 100% of insurers and 50% of pension funds surveyed.

2 Throughout this document EIA refers to economic impact assessment studies commissioned by the Project. These can be found in the Project's webpage: http://www.bis.gov.uk/foresight/our-work/projects/current-projects/computer-trading/working-paper Accessed: 17 September 2012.

3 EIA16 (Annex D refers).

Cliff (EIA16) estimates that it could run to approximately one billion Euros a year if the descriptions were in fact carefully read. Alternatively, the cost of Article 17(2) could be dramatically lowered by simply having firms provide documents to the regulator that are filed but not really analysed. In this case, however, it is hard to see how this activity could actually address the potential risk of algorithmic disruptions to the market.

Care must be taken not to infer that this measure would dramatically reduce systemic risks, and that agents would consequently take larger risks than they otherwise would have. The reason systemic risk may not be reduced significantly even if algorithms were carefully analysed by regulators is that much of the risk arises from the nonlinear interactions of many algorithms. Different algorithms may be present in the markets at different times, setting up what could be infinite combinations of algorithms to consider for regulatory review. Furthermore, even if a 'wind tunnel'[4] for testing algorithmic interaction were constructed, it would not capture all of the systemic risks if algorithms learn and rewrite themselves dynamically over time.

6.1.4 Evidence

There is very little empirical evidence on the cost or benefits of algorithm notification. Cliff (EIA16) provides some estimates of cost, but these depend greatly on how the notification requirements are implemented and on how much the required analysis is split between firms and the regulator. There is to our knowledge no cost estimate of risk controls for algorithms at the firm level.

6.1.5 Conclusions

The desirability of understanding AT strategies and their impact on the market is laudable but achieving this through notification requirements of the type currently envisioned in MiFID II may not be feasible given the complexity of algorithms and their interactions in the market.

6.2 Circuit breakers

6.2.1 The measure and its purpose

Markets have always been subject to episodic price instability, but computerised trading combined with ultra low latency creates increased potential for such instability to occur. This, in turn, has increased interest in the role and usage of circuit breakers. Circuit breakers are mechanisms for limiting or halting trading on exchanges. Their purpose is to reduce the risk of a market collapse induced by a sequence of cascading trades. Such trading could arise from pre-specified, price-linked orders (such as programme trades or trade algorithms that sell more when prices fall), or from self-fulfilling expectations of falling prices inducing further selling. Traditionally, circuit breakers were triggered by large price movements, and hence represented *ex-post* reactions to excessive price volatility in the market. More recently, the advent of HFT taking place at millisecond speeds has resulted in a new generation of circuit breakers which work on an *ex-ante* basis (i.e. halting trading before accumulated orders are executed).

Circuit breakers can take many forms: halting trading in single stocks or entire markets; setting limits on the maximum rises or falls of prices in a trading period (limit up and limit down rules); restrictions on one trading venue or across multiple venues. The London Stock Exchange (LSE), for example, operates a stock-by-stock circuit breaker that, when triggered, switches trading to an auction mode in which an indicative price is continuously posted while orders are accumulated on either side. After some time, the auction is uncrossed and continuous trading resumes. The trigger points are in several bands depending on the capitalisation and price level of the stock, the recent transaction history and the most recent opening prices.

4 What is meant here is a simulated market that can in principle test algorithmic interaction much as a wind tunnel is used for testing aircraft designs.

However, circuit breakers are no panacea. Price discovery is a natural feature of markets, and bad news can induce (sometimes large) price drops to new efficient values. Halting markets can interfere with this natural process, and may simply postpone the inevitable. For example, the October 1987 crash in the USA was subsequently followed around the world, with big price drops in the UK and other markets. The Hong Kong stock market was closed on the Monday after the US markets had started to fall and stayed closed for a week. When it did open, it suffered an instant decline of 30%. On the other hand, the May 2010 Flash Crash in New York was effectively ended by a circuit breaker which allowed liquidity to re-accumulate as buyers returned to the market and the newly balanced market to resume. Thus, circuit breakers, while well suited to dealing with instability caused by temporary shortages of buyers or sellers, are not appropriate for all causes of market volatility and cannot forestall revaluations that are unavoidable.

The policy issue is whether the existing self-regulatory, uncoordinated approach (in Europe) to price instability can be improved.

6.2.2 Benefits

If the price processes are driven purely by rational valuations of fundamentals, then a trading halt impairs the process of valuation and prevents the public from receiving accurate and up to date information. But in today's high frequency electronic markets, liquidity is uncertain, and prices can be affected (at least temporarily) by factors such as imbalances in the book of orders, fat finger trading errors, and errant algorithms engaged in a mechanical feedback loop. Circuit breakers and trading halts may be beneficial for dealing with such imperfections in trading processes. We describe the benefits of circuit breakers in the following three subsections.

Cooling-off period

Many modern markets function as computerised limit order books with continuous trading and replenishment of orders. Even if the daily trading volume is large, the displayed depth of the market at any moment may be relatively small. A large purchase order arriving unexpectedly, for example, can cause a temporary imbalance until more sellers come forward. Circuit breakers provide a respite that prevents mechanical selling at any price, allows the market to understand what is happening and gives counterparties time to enter, thereby reducing the order imbalance.

In a fast-moving market, losses to positions bought using credit can build up quickly, leading the brokers who have provided credit to ask for additional collateral. With very fast-moving markets, these margin calls cannot be satisfied quickly enough and broker confidence may suffer. A cooling-off period allows the traders to raise the collateral, reducing the risk that they fail. It also reduces market risk because brokers will not be forced to liquidate clients' positions which would then put additional selling pressure on the market and create a vicious feedback loop.

Circuit breakers can also be invoked for purely technical reasons to prevent peak overload bottlenecks at the exchanges' servers which could lead to erroneous pricing and execution.

Uncertainty resolution

Volatility is a natural part of markets, but unexplained volatility can cause traders to fear the worst and lead to massive selling. Because high frequency markets move so fast, it may be impossible for traders to evaluate what is causing a large price movement. There is now a large and growing literature showing that uncertainty reduces participation in markets, which can manifest in massive selling for those in the market and a reluctance to participate by those not already there. Either outcome is undesirable, so mechanisms such as circuit breakers that can allow time for uncertainty resolution can be beneficial.

Investor protection

A market that is fast-moving for reasons other than the inflow of fundamental news can not only create systemic risk but can also penalise traditional investors who do not have the resources to monitor markets continuously. Trading halts in response to non-fundamental swings may offer a means of preventing uninformed retail investors losing out to traders who continuously monitor markets. This could bolster the confidence of investors in the integrity of markets, and remove or ameliorate concerns that small investors can be taken advantage of by manipulative trend-generating strategies[5]. A final point is that circuit breakers enjoy widespread support from industry as shown in the survey SR1. They are seen as a prudent mechanism to enforce orderly, fair and efficient markets.

6.2.3 Costs and risks

The obvious cost is that during a market halt traders are prevented from completing mutually beneficial trades. An empirically documented effect of circuit breakers is the so-called 'magnet effect' whereby traders rush to carry out trades when a halt becomes imminent, accelerating the price change process and forcing trading to be halted sooner or moving a price more than it otherwise would have moved. Subrahmanyam (EIA4)[6] recommends that there should be some unpredictability about when circuit breakers are triggered, a logic that may explain Deutsche Börse's decision not to publish the trigger points for its circuit breakers.

Similarly, a trading halt that slows down the fundamental price discovery process may create additional uncertainty. This can undermine confidence and increase bid-ask spreads when markets reopen if the reason for the halt is not credible. More generally, if circuit breakers are implemented across trading venues and coordinated so that a trading halt for a stock on one venue triggers a trading halt for that stock on all other venues, then a halt on a minor trading facility can trigger an unexplained halt on the main venue. In fact, experimental evidence suggests that completely random halts can create uncertainty because traders have the time to make up rumours or to exaggerate existing ones[7].

An additional detrimental element associated with circuit breakers is the inability for market makers to offload large positions quickly when trading is halted. This would have to be factored into their risk management system and might make them less willing to buy, reducing liquidity.

Similarly, there is an issue with respect to cross-asset, cross-venue trade. Suppose a trader is hedging a derivative or is engaged in arbitrage across securities. If trading in the stock is halted, the trader is suddenly no longer hedged and may suffer losses as a result. If this happens frequently, different markets may become less integrated, market efficiency suffers, and the possible loss of confidence in pricing accuracy can lead to feedback loops. If investors suddenly lose confidence in the price of a security (because trading has stopped) and they are invested in other assets in, perhaps, a carefully designed portfolio, they may decide to sell many other assets because their ability to control the total risk they face is compromised. This may lead to a chain of negative events across many securities.

A mandatory, market-wide circuit breaker has not existed in the UK, although in extreme circumstances the LSE declares a 'fast market' and changes certain trading rules (for example, widening individual stock circuit breakers and relaxing market maker obligations). Due to the nature of HFT, much more trading now involves the futures market, exchange traded funds (ETFs), contracts for difference (CFD) and spread betting[8]. The US Flash Crash was triggered in the S&P500 E-Mini futures market, then went to the ETFs on the index, and finally affected the equity market itself.

5 Kim & Park (2010).

6 EIA4 (Annex D refers).

7 EIA9 (Annex D refers).

8 In the UK, CFDs are often used as an indirect way of trading equities but avoiding the stamp duty on equity trading, since CFDs are not covered by this tax.

This raises the problem of how to implement circuit breakers across market venues. In the US Flash Crash, attempts by the NYSE to slow trading were completely ineffective due to the ability of traders to use other venues that were not affected. For a uniform circuit breaker to be effective, it would have to close all markets for a single stock or series of stocks. This would require coordination between and across exchanges, but different exchanges have different trading rules and different trading practices. Moreover, the ability of an exchange to determine its own rules for handling volatility can be viewed as an important dimension of its risk management practices. Regulators would have to consider whether it is desirable for an erroneous trade on, for example, Chi-X to shut down the LSE, or whether the current system whereby only the primary market determines shutdowns is preferable[9].

A hybrid system allowing regulatory override of individual exchange trading halt decisions might provide a mechanism to deal with market-wide disturbances.

6.2.4 Evidence

There is a sizeable literature on the impact of circuit breakers on markets, but overall the empirical results are mixed. Many authors find negative effects of circuit breakers, while others find no effects or small positive effects. However, it is difficult to analyse what would have happened had the market not been halted, and so with a few exceptions these findings are not statistically robust. A more important problem is that high frequency markets are very different from the markets analysed in previous research. There is, as yet, little academic research on the role of circuit breakers in high frequency market settings. In particular, many academic studies are on smaller international markets and are largely concerned with a simple type of breaker rather than the more complex models employed for instance by the LSE[10]. Furthermore, most analyses have almost exclusively focused on transaction prices in the immediately affected market rather than the order book or the spill-over effects on other securities and trading venues.

There is some evidence from the US Flash Crash on the effectiveness of modern circuit breakers. The end of the crash is generally attributed to the imposition of the Chicago Mercantile Exchange's 'Smart Logic', a circuit breaker that halts trading when the accumulated imbalance of pending orders, if executed, would result in the price falling beyond a pre-specified limit. This forward-looking circuit breaker differs from the variety generally employed in most markets which deal with issues once they arise, and may provide a template for the design of market-wide circuit breakers.

There is also some evidence on the use of single stock circuit breakers. The LSE describes how on an average day there are 30–40 trading suspensions, whereas in the first two weeks of August 2011 (when there was a great deal of volatility), this shot up to about 170 suspensions per day. Despite the high number of suspensions, large volume and wide market swings, trading was generally 'orderly' in their view[11]. Of course, there are other market-wide costs associated with each stoppage and it is not clear whether these costs are fully taken into account and appropriately balanced against the benefits.

6.2.5 Conclusions

Circuit breakers have a role to play in high frequency markets, and they are found in virtually all major exchanges. Because of the inter-connected nature of markets, however, there may be need for coordination across exchanges, and this provides a mandate for regulatory involvement. New types of circuit breakers which are triggered before problems emerge rather than afterwards may be particularly effective in dealing with periodic illiquidity.

9 The costs and benefits of coordinated trading halts are further analysed in EIA20 (Annex D refers).

10 A notable exception is a study by Abad & Pascual (2007) who investigate the type of circuit breaker used on the London Stock Exchange and the Spanish Stock Exchange.

11 The London Stock Exchange Group (2011).

6.3 Minimum tick sizes

6.3.1 The measure and its purpose

The minimum tick size is the smallest allowable increment between quoted prices in a market. Tick sizes have important implications for both transaction costs and liquidity provision. The transaction cost effect is straightforward: a larger tick size, if binding, increases trading costs by widening the spread between bid and offer prices. The liquidity effect arises because the tick size determines how easy it is for another trader to 'step ahead' of an existing limit order. In markets with a standard price/time priority (PTP) rule, an order placed first executes ahead of one placed later unless the later order is posted at a 'better' price. Small tick sizes make it easier for traders to post that better price, so smaller tick sizes push up the cost for traders who put trades on the market and provide liquidity. The challenge for markets and regulators is to choose the optimal tick size to balance these liquidity and transaction cost effects. An additional challenge is to decide whether or not minimum tick sizes need to be harmonised across linked venues.

There are important differences in minimum tick size policy between the USA and Europe. In the USA, Regulation National Market System (NMS) requires that in all 'lit' venues (exchanges and large Alternative Trading Systems (ATS)) stocks over $1 are quoted with a minimum tick size of one cent, and sub-penny pricing is prohibited. In Europe, there is no mandated tick size and local exchanges are free to set their own tick policy. As a result, there is generally a range of tick sizes depending on the price level and, in the case of LSE, on the market capitalisation. A European stock may trade on different public venues under different tick size regimes, whereas in the USA, such differences can only currently happen in dark venues. Historically, the trend has been towards smaller tick sizes since US trading in 'eighths' (12.5 cents) yielded to decimalisation in 2000. Now, active stocks in the USA typically trade at one cent spreads, leading to concerns that a one cent minimum tick may be too large thus illustrating the earlier mentioned trade-off between transaction costs and liquidity provision. In Europe, spreads at minimum levels are not as common, suggesting that the tick rules are not as binding on market behaviour. The policy issue is whether it would benefit market quality to mandate a minimum. A related policy issue is whether minimum tick sizes need to be harmonised across venues.

6.3.2 Benefits

Choosing an optimal minimum tick size for a given stock and a given market environment would have a number of benefits[12]. Originally, a uniform tick size rule was chosen to minimise transaction costs for firms and traders. With a limited number of possible prices, the technological complexity of trading was reduced and also the cost. While this is no longer needed given the state of trading technology, it is still the case that a well-chosen minimum tick size framework can reduce the need for firms to split or reverse split their stock in order to influence the relative tick size. This latter reason speaks to the benefits of a non-uniform tick policy as found in Europe.

A well-chosen minimum tick size can prevent market operators such as exchanges, market makers or high frequency traders making excessive profits at the expense of the final users. Chordia (EIA6) argues that a too high minimum tick in the USA led to payment for order flow and internalisation as competitive responses to the excessive profits accruing to liquidity providers in large stocks. To the extent that trading venues compete with each other in a world with HFT and maker-taker pricing [13], exchanges may have an incentive to try to undercut each other's tick sizes to attract volume and hence fees. A coherent overall minimum tick size policy (such as the one agreed to by Federation of European Securities Exchanges (FESE) members, discussed below) that applies to all trading venues could prevent a welfare-destroying race to the bottom where competitive pressures would otherwise lead to a tick size below the optimal range.

12 For further discussion on the optimal tick size and its modelling see EIA7 (Annex D refers).

13 Maker-taker pricing refers to the practice in many exchanges and trading platforms of paying a small rebate to executed orders that were placed as passive limit orders (the liquidity makers) and charging a fee to active orders that hit existing limit orders (the liquidity takers).

A higher minimum limit for tick sizes can also provide greater incentives to post limit orders and thereby create a deeper, more liquid market. Because the costs of jumping the queue are higher with larger minimum spreads, there is less queue-jumping and so a generally more stable book. The reduced frequency of limit order book updating means less data and lower costs related to data. Whether such changes would reduce the arms race of trading firms investing in ever faster trading systems is debatable. A larger tick reduces the incentives to undercut, but it still remains the case that the first trader to post gets the order. This would imply that high frequency traders will generally dominate the order book for active stocks regardless of tick sizes.

Although Europe does not as yet have a mandatory tick size, there have been attempts by the industry through its organisation FESE to harmonise and simplify the tick size regimes across their members. There has been some agreement, but there is no binding legal framework to enforce it and it may not prove sustainable over the long term in the presence of new entrants.

6.3.3 Costs and risks

Setting too large a minimum tick size results in large bid-ask spreads favouring market makers. If prices must be set five pence apart, for example, then the bid-ask spread can also be no smaller than five pence. Except in special cases, a spread of one tick size cannot be undercut because prices must be quoted in increments of the tick size. A large tick size also reduces the chance of a price improvement. A price improvement occurs in any auction when a higher price is offered for the good. If the only bids accepted are in £1,000 increments it is much harder to improve the price than if you are allowed to raise your offer by £1.

An excessive tick size on one market may induce activity to migrate towards venues with lower ticks. In the USA, this is one reason given for the growth of trading in dark venues, and in Europe such tick size competition might be expected to induce orders to flow towards dark or systematic internaliser (SI) venues that are not subject to the rules for multilateral trading facilities (MTF). This migration is most likely to happen with low-priced stocks where the bid-offer spread is large as a proportion of the stock's value. Migration can be mitigated through additional measures, such as coordinated rules across trading venues and SIs, or via mechanisms such as the retail liquidity providers (RLPs) proposed by the NYSE and recently approved by the US Securities and Exchange Commission (SEC)[14]. These RLPs would be allowed to quote on sub-penny ticks provided their quotes are hidden and can only be accessed by retail order flow.

Setting too small a minimum tick size comes with costs as well. In markets with PTP, passive traders submit limit orders and thereby give away a free option to market participants. These passive traders expect to cover these losses through the spread. If tick sizes are very small, then new passive traders can come in and capture the incoming marketable orders by undercutting the current best bid or ask by a tick. Traders who provide liquidity by putting trades on the market will not be rewarded for having taken the risk of being picked off by traders with new information or through adverse selection. Too small a cost for jumping the queue, therefore, makes providing visible liquidity more expensive and leads to smaller and more ephemeral depth. It may also contribute to more cancellations and fleeting limit orders in the book as traders try to snipe in as late as possible.

14 The US SEC approved the NYSE proposed pilot on July 3, 2012 (2012a).

6.3.4 Evidence

There is a large academic literature investigating the influence of tick sizes on market behaviour[15]. In general, the results from studies of a wide range of markets find that a reduction in tick sizes reduces spreads but also reduces depth. As a result, transaction costs for smaller retail investors tend to be lower, but the results are ambiguous for institutional investors whose trades are in sizes that may require greater depth. These empirical findings are consistent with the recent pilot programme implemented by FESE and subsequently analysed by Better Alternative Trading System (BATS) (2009)[16].

Citicorp's recent reverse stock split underscores the effects that changing the relative tick can have on trading[17]. After Citicorp substituted one share for ten, trading in its stock shifted from alternative trading venues to exchanges, HFT activity increased, the trade size rose on alternative venues, volatility was higher and volume lower, and interestingly, order flow toxicity was lower[18]. What is not yet established are the effects on traders or on the issuer.

While the literature carefully delineates the impact of tick size changes on traders and markets, there is no proper analysis on who should make the tick size decision and whether the minimum tick size should be the same for all firms or all trading venues. The current approach in Europe of allowing each venue to choose its own minimum tick size has a variety of merits, but it can lead to excessive competition between venues. The US approach of a uniform minimum tick size removes this problem, but there are deep concerns that it is leading to insufficient liquidity for less frequently traded stocks[19].

6.3.5 Conclusions

Tick size policy can have a large influence on transaction costs, market depth, and the willingness to provide liquidity. The current approach of allowing each European trading venue to choose its own minimum tick size has merits, but can result in unhealthy competition between venues and a race to the bottom. A uniform policy applied across all European trading venues is unlikely to be optimal, but a coherent overall policy for minimum tick size that applies to subsets of trading venues may be desirable. This coordinated policy could be industry-based, such the one agreed to by FESE members.

6.4 Obligations for market makers

6.4.1 The measure and its purpose

Obligations for market makers are requirements that a person (or more controversially, a computer program) acting as a market maker must post prices to buy and sell at competitive levels at all times the venue is open and regardless of market conditions. This could be applied to traditional (human) market makers or to algorithmic market makers or both[20]. The purpose of this proposal is to improve continuous liquidity provision and to ensure that market makers are actively quoting competitive prices during periods of market stress.

Market makers provide liquidity to traders by being willing to buy when a trader wants to sell and to sell when a trader wants to buy. For providing this service, the market maker earns the bid-ask spread. In the case of large, actively traded issues held by both institutions and retail traders, market makers typically earn sufficient returns to justify the time and capital needed to perform this function. For less actively traded stocks, this may not be the case, and wide spreads, reduced depth and general illiquidity may result. In times of market turmoil, market making in any stock is often unprofitable and the withdrawal of market makers can result in market-wide illiquidity. Consequently, market makers

15 See EIA22 for a survey (Annex D refers).

16 See BATS Trading Ltd (2009). London Economics have calculated the economic impact of certain tick size changes through the effect this would have on bid-ask spreads and hence on the cost of capital. See EIA22 (Annex D refers).

17 ITG (2011).

18 Order flow is toxic when it adversely selects market makers, who may be unaware they are providing liquidity at a loss. The fall in order flow toxicity suggests that adverse selection also fell after the reverse stock split.

19 A further issue, as discussed by Angel (EIA7), is that suboptimal tick sizes may lead to stock splits or reverse stock splits by companies who have a different view of what the tick size should be in relation to the price. This splitting phenomenon seems to be more common in the USA, which has a relatively rigid tick size policy.

20 The current MiFID II draft mentions such a proposal in Article 17(3).

have often received various inducements to supply liquidity such as access to fee rebates, or superior access to the order book, or exemption from short sale requirements, or even direct payments from exchanges or issuers. In return for these inducements, exchanges traditionally required market makers to quote bid and ask prices even in times of stress.

In current markets, much of the market maker function has been taken over by HFT in which a computer algorithm is programmed to buy and sell across markets. Such computerised trading often involves the placement of passive orders (i.e. limit orders) and so, like the traditional market maker, the HFT program is buying from active traders who want to sell and selling to active traders who want to buy. However, unlike traditional market makers, the HFT program is not committed to a particular venue, and generally does not have access to superior information, although the computer running the programme may be co-located in the exchange. What is now being considered is whether such high frequency market making should also face obligations with respect to provision of liquidity. Among the various obligations being considered are: maximum spread restrictions; percentage time for quotes to be at the inside spread; minimum quoted size; and minimum quote time.

The policy issue is whether regulators should require market maker obligations for any or all types of market makers. Many exchanges already have some obligations for market makers in terms of quoting, so the question is whether this needs to be mandated across all trading venues or extended more broadly to the market making function[21].

6.4.2 Benefits

Market maker obligations can improve market quality and hence raise social welfare. Narrower spreads will induce both informed and uninformed traders to trade more, which in turn increases price efficiency and quickens price discovery. To the extent that obligations improve the depth of the market, traders will find it easier to buy and sell, and transaction costs should be lower. Obligations to set competitive prices could help reduce volatility, and requirements to stay in the market continuously could lead to greater liquidity during periods of market stress.

It should be noted that benefits from such obligations are not guaranteed because of the high costs that they may entail. When companies have contracted for market making services (as, for example, on the Stockholm Stock Exchange), Weaver (EIA8)[22] reports increased market quality in terms of trading volume and liquidity measures for their shares. However, the high costs of paying for contracting have deterred many small companies from doing so.

6.4.3 Costs and risks

Market maker obligations would be unnecessary if providing liquidity was a profitable and relatively riskless undertaking. The reality, however, is far different, and market makers face a variety of situations in which posting quotes exposes them to the risk of large losses. Moreover, even in the best of circumstances, market making is not cost free, requiring both capital and expensive investments in technology to support operations. To the extent that market making obligations are imposed without corresponding compensation, at least some market makers will exit the market, reducing its liquidity.

Imposing these obligations is problematic. Rules requiring market makers to post narrow bid-offer spreads are often unnecessary for large, active stocks where market making is profitable. However, with less actively traded, small stocks, the order flow is more naturally unbalanced and the market maker faces substantial risk acting as the intermediary for the bulk of trading. Market maker obligations on such stocks will impose significant costs on market makers, and too little market making will be provided unless they are compensated. Conversely, if market makers are given too much compensation, trading in small stocks will essentially be subsidised. Such a situation characterised trading on the NYSE (before the regulatory changes in the late 1990s) whereby small stocks generally benefited from rules restricting maximum spreads and mandating continuous prices.

---

21 For example, NASDAQ has recently proposed the designation of market makers who are compensated by issuers when committing to a minimum liquidity supply. This new programme is under review. US Securities & Exchange Commission (2012b).

22 EIA8 (Annex D refers).

Market making during times of market stress is also an extremely risky proposition as requirements to buy when prices are crashing may lead to bankruptcy for the market maker. An optimal market maker obligation should not force a market maker into bankruptcy, and although limits as to what is actually required are necessary, they are difficult to even define, let alone enforce. Obligations which are too stringent will transfer losses from traders to market makers, while requirements which are too lax can result in greater market instability.

Imposing market maker obligations on algorithmic market making trading strategies raises a variety of risks. Many high frequency strategies post bids and offers across correlated contracts. Thus, a high frequency market maker may be buying in one market and selling in another. A requirement to post a continuous bid-offer spread is not consistent with this strategy and, if binding, could force high frequency traders out of the business of liquidity provision. With upwards of 50% of liquidity coming from high frequency traders, this could be disastrous. A more likely outcome, however, is that any requirement would be evaded by posting one quote at the market and the other off the market or for small size[23]. Moreover, what constitutes market making in this context is unclear. Many high frequency strategies are actually a form of statistical arbitrage, buying where prices are low and selling where prices are high. Using limit orders to implement these strategies is akin to market making, but it also differs in a variety of ways. Forcing market maker obligations on algorithmic trading could reduce these statistical arbitrage activities and thereby reduce market efficiency. These issues are discussed in more detail in Cliff (EIA19).

Cliff (EIA19) also discusses the challenges of specifying a requirement that would apply to algorithmic-based market making strategies in a logical or enforceable way. Changes to the original MiFID II 17(3) specification (known as the Ferber amendments) mean that market maker obligations would apply only to HFT systems that operate on maker-taker trading venues and for which more than 50%, or a majority, of the system orders/trades qualify for maker discount/rebates. While this revised application may be more feasible to implement, it may also induce trade to move to venues where this regulation is not binding.

6.4.4 Evidence

The vast majority of empirical studies on voluntary[24] market maker obligations conclude that they improve market quality. These benefits are found in a wide variety of market settings, and across different classes of securities such as equities and options. The benefits are especially felt in illiquid stocks, where generating critical mass in a market is an issue. The fact that virtually every major stock exchange has some form of market maker obligation testifies to their usefulness in enhancing market behaviour[25].

However, empirical research finds that traders under market maker obligations generate these benefits in part because they get compensated. This can be in the form of extra rights such as sole view of the limit order book, ability to short shares 'naked' (without holding the stock) or direct compensation paid either by the trading venue or by the listing company. Moreover, there is a surprising diversity of approaches taken towards how these market maker obligations are structured. EIA8 notes that "the application of minimum obligations for market makers, as well as the mode of compensation, is uneven across markets on both sides of the Atlantic. Some markets impose minimum obligations on market makers for all listed stocks."

6.4.5 Conclusions

The current system of exchanges determining how to structure market maker obligations and pay for them seems to be working well for most markets. We think there is less support for policies that impose market maker obligations for a large class of market participants without a thought-through incentive scheme to support it.

23 See EIA22 for further discussion of monitoring and enforcement issues (Annex D refers).

24 By 'voluntary' we mean that these all refer to situations where there is some quid pro quo for the market making obligation in terms of fee rebates and superior access to the order flow. Compulsory market making obligations (i.e. without compensation) have not been studied to our knowledge.

25 Of the major markets, only the Tokyo Stock Exchange does not have some form of specified market making. See Weaver (EIA8) for a detailed discussion of these issues.

6.5 Minimum resting times

6.5.1 The measure and its purpose

Minimum resting times specify a minimum time that a limit order must remain in force. The impetus for imposing a minimum is that markets now feature a large number of fleeting orders that are cancelled very soon after submission. This increases the costs of monitoring the market for all participants, and reduces the predictability of a trade's execution quality since the quotes displayed may have been cancelled by the time the market order hits the resting orders. The nature of high frequency trading across markets, as well as the wide-spread usage of hidden orders on exchanges, are responsible for some of this fleeting order behaviour. However, frequent cancelling of quotes may also result from abusive strategies including spoofing, layering, and quote stuffing which can undermine market quality or, at the least, create a bad public perception.

As a result, minimum resting times have been suggested whereby a limit order submitted cannot be cancelled within a given span of time. This measure can take a multitude of forms, such as a uniform 500 milli-seconds across all assets and securities, or a delay that depends on the security and/or general market conditions. It would also be possible to prescribe different minimum resting times on limit orders to buy or to sell, or that adjust to reflect volatility or other market conditions.

6.5.2 Benefits

Minimum resting times can increase the likelihood of a viewed quote being available to trade. This has two important benefits. First, it provides the market with a better estimate of the current market price, something which 'flickering quotes' caused by excessive order cancellations obfuscates. Secondly, its visible depth at the front of the book should be more aligned with the actual depth. This knowledge of the depth improves the ability of traders to gauge the price impact of potential trades. Quotes left further away from the current best bid or offer are less likely to be affected by the measure since the likelihood of them being executed within a short time is small. Nonetheless, minimum resting times might be expected to make the order book dynamics more transparent to the market.

Minimum resting times may also reduce the excessive level of message traffic currently found in electronic markets. Cancellations and resubmissions are a large portion of these messages, and at peak times they can overwhelm the technological capabilities of markets (as seen for example in the recent Facebook initial public offering (IPO) problems on NASDAQ)[26]. Some authors also suggest that minimum resting times may reduce the profitability and incidence of spoofing, quote stuffing and other illicit practices. While conceptually possible, there is no clear evidence that such market abuses only involve flickering quotes, and for those that do, surveillance and fines may prove a more efficient deterrent than imposing a minimum resting time.

Minimum resting times may also allay concerns that markets are currently 'unfair' in that high frequency traders are able to dominate trading by operating at speeds unavailable to other traders. This notion of 'slowing down' markets is not generally supported by economic analyses, but it does speak to the challenge of inducing participation if some traders, particularly small retail investors, feel that speed makes markets unfair.

6.5.3 Costs and risks

Liquidity providers post limit orders available for trade within a period of time in return for an expected gain in the form of the bid-ask spread. Providing limit orders is costly since posting a limit order offers a free option to the market which is exercised at the discretion of the active trader. If an active trader has better or newer information, the limit order poster will be adversely selected, buying when the stock is going down and selling when the stock is going up. As with any option, its value increases with time to maturity and with volatility. Thus, forcing a limit order to be in force longer gives a more valuable option to the active trader, and consequently raises the cost of being a limit order

26 For discussion on the Facebook IPO, see Nanex (2012).

provider. The expected result would be an increase in the bid-offer spread or decreased depth, as posting limit orders will be less attractive[27].

This reluctance to post limit orders will be particularly acute during times of high volatility when the cost of posting the option is naturally increased. This has the undesirable implication that liquidity provision will be impeded just at the times when markets need it most. It also suggests that there could be a feedback effect if increasing volatility triggers orders, further increasing volatility.

A minimum resting time policy may also change the dynamics of the market by attracting more aggressive high frequency traders whose sole aim is to take advantage of the free options. Depending on the length of compulsory resting, those limit orders close to the best bid or offer are likely to become stale (that is, no longer at the efficient price) before they can be cancelled. This can spawn 'front running' by automated traders who collect the low-hanging fruit from such options. In return, the providers of passive quotes will protect themselves against staleness through yet larger bid-ask spreads, or by simply not posting quotes at all. Using the estimates by Farmer and Skouras[28], the cost of hitting such stale quotes may be as high as €1.33 billion per year in Europe[29].

A final argument pointing to larger spreads concerns the nature of market making in high frequency markets. Modern market makers using HFT have, to some extent, replaced capital and inventory capacity by speed. With minimum resting times raising the risk of losses from limit orders, high frequency traders may reduce their market making activities and possibly be replaced by institutional market makers, such as banks. Reduced competition among market makers and their need to earn a return on their capital may drive up transaction costs for end users. Moreover, to the extent that minimum resting times inhibit arbitrage between markets, which is essentially at the heart of many HFT strategies, the efficiency of price determination may be diminished.

6.5.4 Evidence

The empirical evidence to date is very limited, since there are very few natural experiments to shed light on the costs and benefits of minimum resting times[30]. In June 2009, ICAP introduced a minimum quote lifespan (MQL) on its electronic broking services (EBS) platform. These quote requirements set a minimum life of 250 milliseconds (ms) for their five 'majors' (generally currency contracts) and 1,500ms in selected precious metals contracts. In public statements, ICAP credits the absence of a major Flash Crash to MQLs, but of course it is difficult to know what would have happened in their absence. To our knowledge, there has been no empirical analysis of the effects of these MQLs on market or trader behaviour.

Until mid-2011, the Istanbul Stock Exchange (ISE) did not allow the cancellation of limit orders during continuous auction mode unless it was the very last order entered into the system. Depending on the structure and participation on the ISE around the switch, there may be some evidence that can be gathered from that event. Academic research is seeking to identify the effects of this rule change, but the results of any study would be vulnerable to criticism that they could be due to some specific features of the ISE. Clearly, this market is rather different from the LSE or Deutsche Börse, and it is not clear what can be learned from this experiment.

6.5.5 Conclusions

The independent academic authors who have submitted studies are unanimously doubtful that minimum resting times would be a step in the right direction, in large part because such requirements favour aggressive traders over passive traders and so are likely to diminish liquidity provision.

27 EIA21 describes how a minimum resting time would affect different HFT strategies such as market making, statistical arbitrage, 'pinging' (sending and typically immediately cancelling an order to see if hidden liquidity exists) and directional strategies. They suggest a similar impact, although the magnitude of any effect would depend on the length of time that is imposed, market conditions, and the liquidity of the instrument in question (Annex D refers).

28 EIA2 (Annex D refers).

29 These projected losses borne by market makers assume that market makers do not adjust their limit orders, and that aggressive traders can wait until the end of the resting time before they hit the limit order. Competition between latency traders would probably reduce this time, and therefore the resulting profit, by a non-negligible extent.

30 The commissioned study EIA3 (Annex D refers), examined the effects of minimum resting times inside a simulated market. They did not recommend its adoption.

6.6 Order-to-execution ratios

6.6.1 The measure and its purpose

This measure puts an upper limit on the ratio of orders to executions, and as such is part of the larger class of restrictions on order book activity being considered by policymakers on both sides of the Atlantic. The idea of such restrictions is to encourage traders to cancel fewer orders, and thereby provide a more predictable limit order book. It is hoped that such predictability will improve investor confidence in the market. As cancellations and resubmissions form the bulk of market message traffic, this proposal would also reduce traffic and the consequent need for market participants to provide increasing message capacity in their trading systems.

A number of exchanges have some restrictions on messages or the order-to-trade ratio (OTR). The LSE Millennium system, for example, has message-throttling constraints which limit the total number of messages that can come down a registered user's pipes over a 30-second timeframe. It also has a message pricing system which penalises excessive ordering strategies. Sponsoring firms are required to apportion a maximum message rate threshold to prevent sponsored users from entering an overly large number of messages. The limit is set as a maximum number of messages per second per sponsored user and is part of the total limit allowed for the sponsoring firm's allocation. So there are sensible exchange-specific measures already in place that constrain the total message flow and price the externality those messages contribute. The policy question is whether there is value in extra regulation to enforce best practice across exchanges and extend this practice across all trading venues.

6.6.2 Benefits

Receiving, handling and storing messages is costly for exchanges, brokers and regulators. Whenever an economic good is not priced there is a tendency to use more of it than if the user had to pay its actual costs. If the social cost of messages exceeds its private costs, an externality results; the standard solution is to tax messages. A ratio of orders-to-executions (OER) essentially does this, and it can serve to align these private and social costs, thereby reducing the number of economically excessive messages. This, in turn, will reduce the need for exchanges, brokers and other market participants to invest in costly capacity.

With fewer quote cancellations, the order book may be less active and traders may find it easier to ascertain current prices and depths. An OER also may increase the likelihood of a viewed quote being available to trade, partly because passive order submitters would focus more on those limit orders with a higher probability of execution. An OER may also help curtail market manipulation strategies such as quote stuffing, spoofing and layering. Quote stuffing is when a trader sends massive numbers of quotes and immediate cancellations, with the intention of slowing down the ability of others to access trading opportunities. Layering refers to entering hidden orders on one side of the book (for example, a sell) and simultaneously submitting visible orders on the other side of the book (for example, buys). The visible buy orders are intended only to encourage others in the market to believe there is strong price pressure on one side, thereby moving prices up. If this occurs, the hidden sell order executes, and the trader then cancels the visible orders. Similarly, spoofing involves submitting and immediately cancelling limit orders in an attempt to lure traders to raise their own limits, again for the purpose of trading at an artificially inflated price. These strategies are illegal (Trillium Trading in the USA was recently fined by the Financial Industry Regulatory Authority (FINRA) for layering) but they are often hard to detect. By limiting order cancellations, an order-to-execution ratio will reduce the ability to implement these strategies.

6.6.3 Costs and risks

The nature of HFT and market making in fragmented markets naturally implies order cancellations. Algorithmic trading, for example, seeks to reduce trade execution costs by splitting large orders into smaller pieces and sending orders both spatially and temporally to markets. As orders execute or languish, the execution strategy recalibrates, leading to cancellations and resubmissions. Such a trading approach reduces execution costs for traders and leads to greater efficiency in execution. Many HFT strategies (including HFT market making) involve statistical arbitrage across markets whereby movements in a price in one market trigger orders sent to other markets. Again, subsequent price movements in any of the markets will trigger cancellations and resubmissions as part of the process of reducing price discrepancies and enhancing market efficiency.

Many order cancellations are a result of searching for hidden liquidity on limit order books. Exchanges increasingly allow submitted orders to be completely hidden, meaning that the 'best' quotes visible on the book are not actually the best quotes available in the market. To find this liquidity, traders often 'ping' or send small orders inside the spread to see if there is hidden liquidity. Because such orders are typically cancelled, a binding OTR would result in less pinging and, therefore, less information extraction at the touch. As a result, more hidden orders will be posted, leading to a less transparent limit order book. A second effect on the book may arise because orders placed away from the touch (the best bid and ask prices) have the lowest probability of execution. In a constrained world, these orders may not get placed, meaning that depth may be removed from the book away from the touch. An added difficulty is that the constraint may be more likely to be binding during times of extreme market activity. Brogaard (EIA1)[31] argues that this will reduce the willingness of traders to post limit orders during volatile times, thus reducing market liquidity provision when it is most needed.

Finally, there is the calibration exercise of where exactly to set any ratio and to what type of orders or traders it will apply. If the upper limit of the OER is small, then it will stifle legitimate activities and prevent socially useful trading. For instance, ETFs and derivatives valuations may become unaligned, leading to inefficient pricing. Because of this, the London Stock Exchange (LSE) has an OTR of 500/1 for both equities, ETFs and exchange traded products (ETPs), with a high usage surcharge of five pence for equities and 1.25 pence for ETFs/ETPs. If instead the upper limit is set high enough not to impinge on legitimate order strategies, it may not have much impact on the market either (a point made by Farmer and Skourous (EIA2)[32]). If the intent is to limit manipulative strategies, a specific charge for messages (and greater surveillance) may be a better solution[33].

6.6.4 Evidence

There have been no published academic studies of OERs, and this greatly limits the ability to gauge the costs and benefits of order activity restrictions in general, and OERs in particular. The commissioned study, EIA18[34], investigates the effect of the introduction of an OER penalty regime on the Milan Borsa on 2 April 2012. The authors' preliminary findings are that liquidity (spreads and depth) worsened as a result of this policy measure. They also find that the effect is more pronounced in large stocks, although they acknowledge some issues with their methodology.

There are a variety of actual market programs that provide some evidence of OER impact. The ICAP has a monthly fill ratio (MFR) requiring that at least 10% of all quotes submitted into the market must result in an execution. Similarly, LSE's Millennium trading system has message throttling constraints and penalties for excessive ordering strategies. Anecdotal evidence suggests that the LSE message policy was not fully effective in that it gave rise to new patterns of trade in low-priced stocks. The LSE has experimented with changes in pricing effective May 4, 2010 whereby, among other measures, the threshold for the high usage surcharge for FTSE 350 securities increased from an OTR of 100/1 to a ratio of 500/1 (which is still the figure in use at the time of writing). The frequency of order book updates nearly doubled for a few months as a result before coming down again. Unfortunately, we are not aware of a proper scientific investigation of these effects.

6.6.5 Conclusions

An OER is a blunt measure that constrains both abusive and beneficial strategies. It may not do too much harm if the upper limit is large enough not to hinder market making and intermediation, but to the extent that it is binding on those activities it may be detrimental to both spreads and liquidity. It is unlikely that a uniform OER across markets would be optimal because it depends upon the type of securities traded and the trader clientele in the market. If a ratio could be structured to target

31 EIA1 (Annex D refers).

32 EIA2 (Annex D refers).

33 EIA21 outlines a number of ways in which an OER could be bypassed or manipulated by HFT, whereby any benefits from the rule may be reduced (Annex D refers).

34 EIA18 (Annex D refers).

those quotes that are considered socially detrimental directly, then it might be a useful tool for combating market manipulation. The absence of research which investigates costs and benefits, as well as the difficulty of actually setting this measure optimally, suggest caution in adopting this approach for the market.

6.7 Maker-taker pricing

6.7.1 The measure and its purpose

Maker-taker pricing refers to a fee structure in electronic markets whereby providers of liquidity via limit order submissions (or 'makers') receive a rebate on their executed orders while active traders (or 'takers') pay a fee for executing against these limit orders. The current fee structure at the LSE (and many other European exchanges) uses maker-taker pricing and features both a quantity discount and a rebate for suppliers of liquidity.

The issues with maker-taker pricing are twofold. First, does this pricing scheme incentivise high frequency traders to dominate the provision of liquidity and exclude other traders from posting limit orders? Second, does it raises agency issues for brokers who route orders based on the net fees and who may as a result not provide best execution for their clients?

The policy issue is whether to limit maker-taker pricing with the stated objectives of reducing the role of HFT in markets and making order-routing decisions more transparent.

6.7.2 Benefits

In electronic markets, liquidity is provided by limit orders that are posted by passive traders willing to provide an option to active traders. The more limit orders posted on a market, the more liquidity there is for other traders to execute against. If there are no frictions in the market, then the bid-ask spread would settle at a level that exactly compensates the providers of liquidity with the value received by takers of liquidity (see Foucault (EIA12)).

Real markets do have frictions, however, so fee and pricing models are not irrelevant. By paying those who make liquidity while charging those who take liquidity, maker-taker pricing has the potential to improve the allocation of the economic benefits of liquidity production. This in turn can incentivise potential suppliers of liquidity and lead to faster replenishment of the limit order book. Varying maker-taker fees with market conditions also provides a means to improve liquidity provision during times of market stress. A recommendation to have such time-varying fee structures was one finding of the Commodity Futures Trading Commision (CFTC)-SEC's Task Force On Emerging Market Issues (the Flash Crash commission).

Maker-taker pricing can also be an effective way for new market venues to compete against established venues. Smart order-routing systems can direct order flow to venues with more aggressive pricing models. That can in turn put pressure on fees in other markets and lead to more competitive pricing. The success of BATS in the USA is often attributed to their aggressive use of maker-taker pricing. An interesting competitive development in the USA has been the arrival of trading venues offering 'taker-maker' pricing. In these venues, providers of active order flow get rebates while providers of passive order flow face charges. Venues offering this pricing model are attempting to attract the 'less toxic' orders of retail traders, and market makers pay for the privilege of interacting with this order flow. In US options markets, both models exist simultaneously, suggesting that a variety of pricing models may be viable in heterogeneous markets[35].

6.7.3 Costs and risks

High frequency traders are generally better able to put their limit orders at the top of the queue due to their speed advantage and to their use of 'big data' to forecast market movements. The maker-taker fee structure may incentivise them to do so even more, with the result that institutional investors' limit orders are executed only if high frequency traders find it uneconomical to do so. It follows that institutional investors will hold back from submitting limit orders, leaving the market vulnerable to

35 Anand *et al.* (2011).

transient participation by high frequency traders during times of market stress. This, in turn, could exacerbate episodes of periodic illiquidity.

It also follows that because institutional investors will now submit more market orders, they will face increased costs arising from bid-ask spreads, and taker fees. This problem of higher trading costs can be compounded if the routing decisions taken by intermediaries on behalf of clients are influenced in a suboptimal way by the fee structure offered by disparate venues. In particular, the broker may opt to send orders to venues offering suboptimal execution in return for rebates that are not passed on to the originating investor. This incentive will be even greater if these rebates are volume-dependent. Because European best execution requirements are complex, it may not be easy to monitor such practices.

A complex system of maker-taker pricing that is context- and venue-dependent can confuse market participants and lead to erroneous decisions. This may be particularly true if markets vary fees and rebates across time. Because spreads can vary, it is not entirely clear how much incremental effect on liquidity will arise from time-varying rebates.

6.7.4 Evidence

There is little evidence in the academic literature that high frequency traders are 'abusing' the current fee structure. Hendershott and Riordan (2011) find evidence that high frequency market makers lose money in the absence of rebates[36]. This is consistent with competitive pricing strategies forcing spreads to lower levels than they would be in the absence of maker-taker pricing.

Examining the effects of a controlled experiment on maker-taker pricing on the Toronto Stock Exchange, Malinova and Park (2011) find that the bid-ask spread adjusted to reflect the breakdown of maker-taker fees[37]. They also found that the quoted depth of stocks eligible for maker-taker pricing increased significantly, suggesting provision of greater liquidity. Adjusting for the fees, the average bid-ask spread was the same before and after the introduction of maker-taker pricing, and volume was greater for those stocks. Overall, maker-taker fees improve markets by increasing depth and volume while holding spreads (including fees) the same. Anand *et al.* (2011) compares the make-take structure with the traditional structure in options markets, where exchanges charge market makers and use payments for order flow. They find that neither structure dominates on all dimensions but that the make-take structure is more likely to narrow existing quotes, attracts more informed order flow, draws new liquidity suppliers, and performs better in lower tick size options.

A similar study investigated the introduction of maker-taker exchange fees for Australian securities cross-listed on the New Zealand Stock Exchange in 2008. Berkman *et al.* (2011)[38] found that depth at the best quotes as well as trading activity increases with the introduction of maker-taker fees, though there is little evidence of a change in bid-ask spreads.

The only study focusing specifically on maker-taker pricing in European markets identified by the Project is Lutat (2010)[39]. He finds that the introduction of maker-taker pricing by the SWX Europe Exchange did not affect spreads but led to an increase in the number of orders at the top of the book. Incidentally, Menkveld (2012) finds that spreads fell dramatically when Chi-X began trading Dutch index stocks, suggesting that its maker-taker model may have improved market competitiveness[40].

6.7.5 Conclusion

Overall, the limited evidence suggests that maker-taker pricing improves depth and trading volume without negatively affecting spreads.

36 Hendershott & Riordan (2011).

37 Malinova & Park (2011).

38 Berkman *et al.* (2011).

39 Lutat (2010).

40 Menkveld (2012).

6.8 Central limit order book

6.8.1 The measure and its purpose

A central limit order book (CLOB) would consolidate all limit orders into a single queue for trading. This, in turn, would reduce the incidence of locked markets (when the best price to buy is the same as the best price to sell) or crossed markets (when the best price to buy is lower than the best price to sell). A central queue would also ensure that limit order providers are treated fairly in terms of price or time priority. An actual centralised single order book is not necessary to achieve a CLOB, although that is one option. Another alternative is a decentralised US-style consolidated quote and trade tape: this can work to create a virtual CLOB when combined with rules specifying 'trade at' (the first order at a price must be executed first) or 'trade-through' (another venue can trade an order provided they match the better price posted on another market) protection. The policy issue is whether Europe should establish some form of a CLOB.

6.8.2 Benefits

In Europe, best execution requirements are not formulated solely in terms of price and so linkages between markets in Europe are not as tight as between markets in the USA. This is demonstrated by the higher frequency of locked and crossed markets and trade-throughs in Europe, where stocks can trade at different prices at the same time on different venues. Moving to a CLOB could reduce such differences and improve market quality.

A CLOB also reduces queue-jumping across limit order books and therefore encourages the display of liquidity. Having a stock trade only in a single market setting, however, represents a step backward to the situation before the first round of MiFID. A virtual CLOB could ease this concern, but it might require extensive regulatory rule-making to set out how the various markets would be linked. The growth of HFT, with the consequent increase in arbitrage across markets, may help reduce these market abnormalities without having to establish a single CLOB. The benefits of a CLOB would be reproduced in a virtual CLOB by enhancing transparency and linkages between markets.

6.8.3 Costs and risks

Recombining local limit order books into a CLOB poses a variety of technological and economic difficulties. Creating a single pool of liquidity confers monopolistic power on a trading venue, which allows it to increase fees and costs which are then passed on to investors. A single physical pool may also stifle competition as smaller venues may be unable to differentiate their services. MiFID was intended, in part, to address the problems caused by the centralisation of trading.

Creating a virtual CLOB as is partially accomplished in the USA raises a host of issues. While trade-throughs and locked and crossed markets are relatively rare in the USA, internalisation and order preferencing[41] have flourished in part due to the access fees required by US routing rules. Moreover, US rules provide only top of book protection, meaning that once the best posted order is executed, the broker or venue is free to execute the rest of the order at inferior prices prevailing on their exchange. This discourages limit order provision as better orders can be passed over if they are at another venue.

A major problem with any centralised limit order implementation is cost. Creating a single physical order book for all of Europe is probably unfeasible. A virtual order book created by the simultaneous linking of all European trading venues requires both instantaneous access for all traders, a central tape to provide information on where the best price is available and complex technology to keep track of shifting order priorities induced by trade executions. The Project was not able to find any cost estimates for building such a system.

41 Order preferencing refers to arrangements whereby orders are sent to a pre-specified exchange or broker/dealer in exchange for a per share payment. When these arrangements are made with a broker/dealer firm, the firm directly crosses these orders with its other order flow, allowing the trade to occur internally (i.e. internalisation).

6.8.4 Evidence

There is a growing literature suggesting that current markets can achieve the consolidation benefits of a CLOB without having to move to a structure for a CLOB. Foucault and Menkveld (2008) show that when EuroSETS introduced a competing limited order book for Dutch equities (up to then, traded only on Euronext) spreads were largely unaffected, controlling for other factors, and consolidated market depth was up substantially by 46% to 100%[42]. However, depth measures must be interpreted with caution given the ephemeral nature of orders in some electronic venues. That is, the larger depth may be posted by the same high frequency traders which may cancel quotes on one venue once the quote on the other venue is hit.

Gresse (DR19) compares European stocks before and after the implementation of MiFID in November 2007 and shows quoted and effective bid-ask spreads are smaller on average after the implementation of MiFID. This evidence suggests that the fragmentation of order flow has not adversely affected European transaction costs. Degryse *et al.* (2011) analyse 52 Dutch stocks (large and mid-cap) over the 2006–2009 period[43]. Moderate levels of market fragmentation (as measured by the Herfindahl index) are associated with an improvement in market liquidity, but too high a level of market fragmentation is harmful for market liquidity, especially for the primary market. Ende and Lutat (2010), Storkenmaier and Wagener (2011), and Gomber *et al.* (2012) discuss the issue of trade-throughs and locked and crossed markets in European markets[44]. They generally find that these market failings are not as bad as some reports have claimed and that they have improved over time, indicating that a virtual market is emerging. For example, in April/May 2009, 85% of quotes for FTSE100 companies coincided with a positive spread across the main UK venues, while in April/May 2010 this percentage had increased to 94%.

6.8.5 Conclusion

The evidence suggests that the benefits to consolidation of trading into a single order book have been almost achieved by the virtual consolidation of fragmented trading venues.

6.9 Internalisation

6.9.1 The measure and its purpose

Internalisation refers to the practice whereby some customer trades are executed internally by brokers or intermediaries and so do not reach public markets. This can provide benefits to both parties of the transaction as well as to the broker who arranges it.

Internalisation is part of a broader set of issues on order routing and pre-trade transparency. This is because in the process of trying to find a counterparty for a client's order, the broker may also route the order to dark pools that are run by groups of brokers rather than to the public limit order book. To address this issue, MiFID introduced a category of trading venue called systematic internaliser (SI) which deals on its own account by executing customer orders in liquid shares outside a regulated market or a MTF, subject to certain transparency requirements. For example, Goldman Sachs and Nomura operate SIs in the UK. Since their introduction, the volume traded on SIs in Europe has increased a little but still represents only a small fraction of the total traded volume. In contrast, the over-the-counter (OTC) category has grown considerably.

There are a variety of concerns connected with broker crossing networks[45]. One is that crossing networks have avoided the burdensome rules associated with the SI category and have kept their liquidity dark and unregulated. There are also worries that internalisation has increased to an extent that it poses a risk to overall market quality. This is because keeping order flow out of the public limit order book might adversely affect the liquidity of public markets and the effectiveness of the price

42 Foucault & Menkveld (2008).

43 Degryse *et al.* (2011).

44 Ende & Lutat (2010) Storkenmaier & Wagener (2011), Gomber *et al.* (2012).

45 Broker crossing networks are trading platforms in which traders submit buy and sell orders and trades are executed for matching orders at an exogenous price, typically the midpoint of the current bid and offer spread. As there are often more orders on one side than the other, not all orders execute. Trades that do execute, however, do so at a better price than they would have in the market.

discovery process. Finally, there are concerns about possible negative effects from internalisation during a crisis period such as the Flash Crash.

The commissioned study EIA10[46] considers several specific policy measures that have been proposed in the USA to limit or constrain internalisation. The trade-at rule would forbid internalisation unless the venue is itself displaying the best bid or offer at the time the order is received, or provides significant price improvement on the order. The rationale for this rule is tied to the current US price-only best execution system, which does not operate in Europe. The sub-penny pricing rule would allow RLPs to quote in minimum tick sizes on public exchanges smaller than one penny, hidden from the view of the wider market, and accessible only to providers of retail order flow. The purpose of this rule would be to reduce the incentive to use internalised venues which can use smaller tick sizes than public exchanges. The minimum size requirement rule would restrict the size of order that an internaliser can execute, to direct more retail flow to the public markets. The dark pool quote threshold rule would require internalisers to display a stock's trading if overall trading of that stock by internalisers exceeded a set proportion, such as 5% or 25%, in order to improve the transparency of trading.

6.9.2 Benefits

Internalisation essentially diverts order flow from public markets to private venues. This has two undesirable effects. First, internalised orders do not contribute to the costly process of price discovery but instead 'free ride' on this function performed by public markets. Exchanges and other public markets argue that this is unfair as they are not compensated for this activity by the internalisers.

Second, to the extent that the diverted order flow comes from uninformed retail traders, it is less 'toxic' than order flow that is not internalised. Order flow is considered 'toxic' when it is related to possible future price movements. Thus an informed trader will be buying when there is good news, hoping to profit from the information when prices subsequently adjust. The market maker, however, is on the other side of the trade, and so generally loses to informed traders. If order flow is too toxic, spreads have to widen and market makers may even curtail the provision of liquidity. Exchanges argue that internalisation, by taking the non-toxic flow, makes the orders coming to the exchanges more toxic and reduces market quality on the exchange.

Reducing internalisation may also enhance competition in the market. In particular, price matching preferencing arrangements eliminate the incentive to compete on the basis of price, so a reduction in internalisation may reduce spreads. Bloomfield and O'Hara (1999) found in an experimental study that order preferencing led to less competitive market outcomes once preferencing in the market reached a threshold level[47].

Restraining internalisation may reduce the potential for some conflicts of interest. For example, it has been alleged that when prices fell during the Flash Crash, internalisers executed buy orders in house but routed sell orders to exchanges, adding further pressure on liquidity in these venues. Furthermore, with US-style order protection rules, orders must be executed at the best bid or offer. But in times of market stress, the consolidated tape may be slower to update than the proprietary data feeds purchased directly from the markets. This sets up the potential for internalisers to give clients the stale national best bid and offer price while simultaneously trading at better prices available in the market.

6.9.3 Costs and risks

In normal circumstances, internalisation is mutually beneficial for participants: brokers execute client orders, avoid paying exchange and clearing fees, and receive non-toxic trades; clients execute their trades without signalling their intentions publicly or incurring market impact. Best execution rules in the USA require that internalised orders are executed at the current best bid or offer. Best execution rules in Europe differ, but internalised trades often achieve better prices and are executed inside the spread. Thus, for many traders, internalisation is beneficial. Requiring a minimum size, however, would effectively preclude retail order flow from being internalised, while a trade-at rule would dramatically

46 EIA10 (Annex D refers).

47 Bloomfield & O'Hara (1999).

reduce the quantity of internalisation. Both changes would be expected to raise transaction costs for retail traders.

If minimum size thresholds are mandated, institutional investors can continue to trade large blocks of shares through internalisers, but the volume coming from retail investors will fall, leading to lower profitability for the internalisers. With a dark pool quote threshold rule, institutional investors with a large block-trading demand may not be able to trade opaquely. This, in turn, could discourage institutional shareholding (especially of small-cap stocks) in the first place.

Finally, restricting internalisation would remove some liquidity providers from the market. This might be expected to increase spreads and transactions costs for all traders.

6.9.4 Evidence

There are a variety of studies that provide evidence on this issue. Weaver (2011) analyses the effect of the increase in internalisation on market quality using NYSE, AMEX and NASDAQ data from October 2010[48]. Controlling for variables known to be associated with spreads, he finds that internalisation (as represented by trades reported to a trade reporting facility (or TRF)) is directly associated with wider spreads. For example, a NYSE listed stock with 40% of TRF volume has a quoted spread $0.0128 wider on average than a similar stock with no TRF reporting. However, institutional investors trading through internalisers may in fact have traded within the spread set on the public exchanges, so these order book findings cannot be used to draw conclusions about the welfare consequences as there is no evidence given on the actual transaction costs faced by these traders. Weaver does find that higher levels of internalisation are always associated with higher levels of return volatility.

O'Hara and Ye (2011), however, find very different results using TRF data[49]. These authors view TRF reported trades as reflecting fragmentation rather than internalisation *per se*. This is because TRF data also include trading on non-exchange venues such as Direct Edge and BATS (both of which have since become stock exchanges). Using a matched sample approach, O'Hara and Ye find that stocks with greater TRF volume had lower spreads and greater informational efficiency. They interpret this as evidence that fragmentation does not reduce market quality, but they refrain from offering conclusions on internalisation.

A third relevant study is by Larrymore and Murphy (2009)[50]. In October 1998, the Toronto Stock Exchange imposed a new rule that effectively banned internalisations of limit orders for 5,000 shares or less unless they resulted in better prices. They find a statistically significant improvement in market quality as a result, with spreads down, quoted depth up and volatility down.

6.9.5 Conclusion

In summary, the data relating to this measure are mixed at this point as to the impact of internalisation on market quality. Some internalisation of trades is necessary and beneficial, but there must be a limit, which if breached would result in negative outcomes for many participants. It is harder to say where precisely markets are currently located on this range of outcomes, so policy makers should err on the side of caution before being over- prescriptive about these practices.

6.10 Order priority rules

6.10.1 The measure and its purpose

Order priority rules determine the sequence in which submitted orders are executed. Most exchanges have migrated to electronic limit order books using price/time priority (PTP). This means that a limit order is queued and executed when there is trade at that price on a first-come first-served basis. However, there are exceptions or qualifications to this rule in practice. Hidden orders generally take a lower priority than visible ones, partially hidden orders are separated into the two queuing

48 Weaver (2011).

49 O'Hara & Ye (2011).

50 Larrymore & Murphy (2009).

priorities and contractual preferencing arrangements may also violate strict PTP[51]. The policy issue is whether PTP unduly rewards high frequency traders and leads to over-investment in an unproductive technology arms race.

6.10.2 Benefits

The greatest benefit of a PTP rule is that it treats every order equally. Using other priorities, such as a *pro rata* rule where every order at a price gets a partial execution, gives greater benefits to large traders over small traders. In addition, PTP provides a stronger incentive to improve the quote than a *pro rata* rule, enhancing liquidity dynamics. Limit order providers face risks in that traders with better information can profit at their expense. PTP encourages risk-taking by giving priority in execution to limit order providers willing to improve their quotes.

However, if PTP is not operating across platforms (which is typically the case in Europe), then the magnitude of the time priority incentive is potentially greatly reduced by a substantial increase in fragmentation. Fragmentation reduces incentives for limit orders to compete aggressively in order to jump the queue.

PTP is also a convenient tool used by traders for directional strategies by, for example, buying through a market order while selling with a limit order and waiting.

6.10.3 Costs and risks

With a PTP rule, the first order in at a price executes first. This priority rule encourages the race for speed, and fosters practices such as collocation with exchanges or investments in costly technology to be earlier in the queue. There are serious concerns that the quest for relative speed leads to over-investment in an unproductive 'arms race'.

Alternative priority rules such as call auctions (for example, TAIFEX for trading the Taiwan Stock Exchange Stock Index Futures) avoid the issue of the speed priority. They may reduce the race for speed, lead to considerable aggregation of information and limited adverse selection when trading arises, and be less prone to (mini) flash crashes. Crossing networks traditionally featured a single clearing price (as is the case with call auctions), although contemporary crossing mechanisms feature a variety of priority rules.

Other mechanisms have been suggested to extract information better, such as spread-priority (a spread order is placed on both sides of the market) ahead of price and time[52]. If a person has private information and knows that a security is underpriced, then he/she does not like to quote both ways. Spread-priority would imply that he/she needs to put in a higher bid than he/she otherwise would have, giving away some information. Of course, if all information is extracted, then the Grossman-Stiglitz paradox obtains[53].

6.10.4 Evidence

There is not much empirical research comparing the benefits of different priority rules. Call auctions are used at the open and close on many exchanges, and so investors can choose to submit orders at those times if they prefer an alternative priority structure.

51 Preferencing arrangements generally promise to execute an order at a given price (typically the current best bid or offer at the time of the trade). These orders violate price/time priority because they can execute before orders place first.

52 CBOE (Chicago Board Options Exchange) Rulebook. Paragraph (e) of Rule 6.45, Priority of Bids and Offers, provides a limited exception from the normal time and price priority rules for spread orders. Spread orders are defined in Exchange. Rule 6.53(d) to include only those spreads that involve the same class of options on both legs of the spread. Rule 24.19 provides an exception from the normal priority rules for an OEX-SPX.

53 The Grossman-Stiglitz paradox states that informed traders will only gather new information if they can profit on its trading, but if the stock price already incorporates all information then there is no way to profit on it. But if no one gathers new information, how did the information get into the stock price? See driver review DR12 (Annex D refers) for a discussion of this paradox and its role in the current high frequency environment.

6.10.5 Conclusion

Without undertaking a compelling cost-benefit analysis, the benefits of rule changes currently appear insufficient relative to the costs. There does not appear to be a strong case to change substantially the system of order priority.

6.11 Periodic call auctions

6.11.1 The measure and its purpose

Most equity trading is now undertaken through the mechanism of electronic limit order books. There are some arguments that this incentivises expenditure designed to achieve a relative speed advantage, and that speed of this order brings no social benefit[54]. An alternative trading mechanism can be based on periodic auctions, which can be designed to minimise the advantage of speed and to mitigate other negative outcomes of the continuous trading model such as manipulative strategies.

The specific proposal in EIA11 involves a sequence of intraday auctions of random starting point and duration that accumulates orders and then executes at a single price using a *pro rata* mechanism to allocate outcomes. There would be a cap on the size of orders to be included and each auction would be a sealed bid, meaning that during the bidding process no information is disseminated to bidders. These intraday auctions would differ from existing auctions employed at the open and close of the business day, or the auctions used to re-start trading after a trading halt.

These call auctions would replace continuous trading and serve to focus the attention of market participants on fewer, centralised trading events. There could also be a hybrid approach in which call auctions are interspersed throughout the day with continuous trading, but as of yet no market has proposed or adopted such an approach.

6.11.2 Benefits

The main benefit of periodic call auctions would be a reduction of the speed of trading and the elimination of the arms race for speed discussed above. Indeed the speed of trading could be controlled through the timing and frequency parameters which could be tuned to individual and market conditions. It might increase liquidity or at least concentrate it at particular time points. It may also decrease the likelihood of short but dramatic deteriorations in liquidity which have been documented by Nanex[55].

6.11.3 Costs and risks

While in theory a sequence of batch auctions can be *Pareto* optimal (improving someone's welfare without reducing another's welfare), little is known about their efficiency in real-world securities markets where the structure of the economy is not common knowledge. Double auctions, on the other hand, have been the dominant financial market mechanism for many years, and their great efficiency in allocating resources has been well documented, leading Vernon Smith famously to describe this finding a 'scientific mystery'[56].

One issue with periodic batch auctions is the increased execution risk that investors will face if they do not know when the auctions will take place and then whether their bid will be in the winning set. If the frequency of auctions does not reflect market conditions, there may be some shortfall in matching supply and demand.

Periodic call auctions would have a severe impact on the business model of market makers and may reduce incentives to supply liquidity. It would also seriously affect hedgers who are trying to maintain positions across equities and derivative markets. The current market structure allows a variety of different trading mechanisms and allows for auction-based trading as well as the continuous limit order book in both lit and dark modes. This proposal would limit choice and require quite a drastic change

54 See for example Haldane (2011).

55 Chapter 4 has more discussion on this. See http://www.nanex.net/flashcrash/ongoingresearch.html Accessed: 17 September 2012.

56 Smith (1982).

in the trading landscape. Furthermore, it would require coordination at the global level, which may be difficult to achieve.

6.11.4 Evidence

Historically, auctions were more prevalent in equity markets than they are today: the Paris Bourse relied on a daily auction until the late 1980s. These mechanisms were gradually replaced by the electronic order book, but auction mechanisms have seen a resurgence. The NYSE, NASDAQ, Deutsche Börse, Euronext and the LSE all have opening and closing auctions as well as continuous trading. These exchanges now also make considerable use of auctions to restart trading after intraday trading halts. Amihud *et al.* (1997) investigate the transition from daily call auction only to daily call auction followed by continuous trading on the Tel Aviv Stock Exchange in 1987[57]. They find improvements in liquidity and market quality attributable to the change, which suggests that continuous trading adds positive value.

Nowadays, auction mechanisms can vary in important directions such as the degree of transparency reported during the bidding period and whether they have a random element to the ending time or not. One study by Bommel and Hoffmann (2011) compares the French Euronext auctions (which report five levels of limit orders and have fixed ending times) with the German Xetra auctions (which only disclose the indicative price and volume and have random ending points)[58]. After controlling for stock characteristics, the evidence shows that the Euronext auctions are more liquid, contribute more to price discovery, and are followed by lower bid-ask spreads in the continuous trading. The more transparent market (Euronext) is superior in terms of price and quantity discovery, and price impact, but both auction mechanisms are followed by price reversals in continuous trading.

The only exchange that apparently makes use of a similar mechanism to the periodic call auction is the Taiwan Stock Exchange Stock Index Futures (except that it reports indicative prices and does not have restrictive volume caps). There does not appear to be any empirical evidence to compare current mechanisms with the proposal discussed here. However, market participants seem to have a wide range of choices about how and where to trade including access to auction mechanisms. Kissell and Lie (2011), for example, document that there have been no long-term trends in auction volumes on the NYSE or NASDAQ[59].

6.11.5 Conclusion

The evolution of many crossing networks (see, for example, POSIT) from discrete crossing (essentially a call market) to continuous crossing, suggests a market preference (at least among institutional traders) against call auctions.

6.12 Key interactions

In this chapter the different measures available to policy makers have been discussed individually. A number of key interactions between the measures are considered in the section below.

The presence or absence of circuit breakers affects almost all the other measures except perhaps notification of algorithms. The direction of the effect is harder to determine. Having more stable and orderly markets is likely to improve conditions for many traders, but decreasing the probability of execution for limit orders may adversely affect particular trading strategies.

Likewise, minimum tick sizes affect almost every other measure except perhaps notification of algorithms. The tick size affects the profitability of market making and so will affect market maker obligations. The smaller the tick size, the more onerous are the obligations to post competitive bid-offer quotes because the return from doing so (the spread) is smaller. Internalisers and dark venues do not currently face minimum tick size restrictions and can use smaller tick sizes to lure trades away

57 Amihud *et al.* (1997).

58 Bommel & Hoffmann (2011).

59 Kissell and Lie (2011).

from the public limit order book. Limiting internalisation and raising tick sizes therefore need to be considered together.

Maker-taker pricing also becomes more powerful the larger the tick size. Therefore, encouraging intelligent maker-taker pricing and larger minimum tick sizes need to be considered at the same time. PTP and minimum tick sizes are also interrelated. Time priority encourages liquidity submission if the queue cannot be jumped at low cost, and the tick is that cost. The tick size affects the profitability of market making and so will affect market maker obligations. The smaller the tick size, the more onerous are regulations to post competitive two-sided quotes because the return from doing so (the spread) is smaller.

Minimum resting times and minimum tick sizes may complement each other on passive orders and possibly conflict on active orders. One of the assumed benefits of minimum resting times is a slowing down of (passive) activity. Larger tick sizes have this effect as they discourage queue-jumping and increase the value of being towards the front of the queue. Larger tick sizes make speed more valuable as it improves the chances to be placed towards the front of the queue, but minimum resting times make this more dangerous for the trader. In that sense, the measures are complementary since minimum resting times blunt to some extent the speed advantage granted by larger minimum tick sizes to faster limit order traders. But minimum resting times also make speed for market orders more valuable as the fastest aggressive order will be first in picking off a now stale passive order. If ticks are larger, this opportunity will be more profitable still, albeit rarer. The interaction is therefore complex and ambiguous.

If minimum resting times and lower minimum tick sizes are introduced for a security, the benefits of either measure may not be fully reaped. The reason is that the existence of a minimum resting time by itself tends to increase bid-offer spreads on one hand, and on the other hand as the true value of the security moves (*ceteris paribus* given the minimum tick size), it is more likely to render resting quotes stale if tick sizes become smaller. This makes 'picking off' more frequent (but less profitable). It follows that given minimum resting times, a reduction in minimum tick size may not lead to a significant reduction in spreads as passive order submitters need to protect themselves against more frequent sniping.

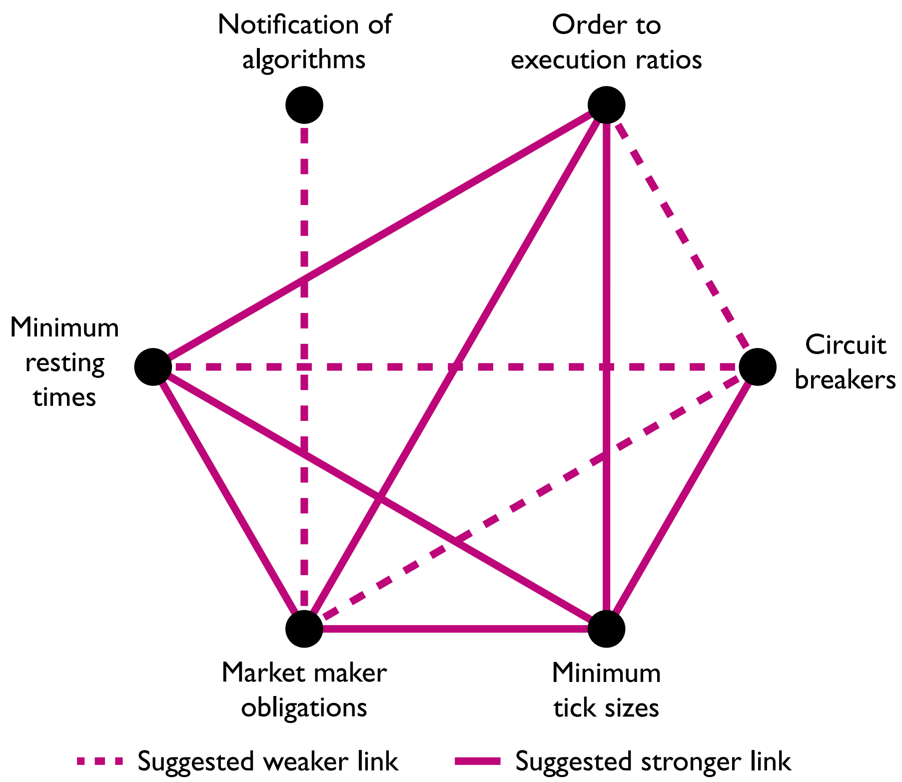
OERs and larger minimum tick size both reduce traffic and complement each other. Depending on the nonlinearities between quote volume, server speed and stability, quote stuffing may become easier given already large volumes of data, in which case a larger minimum tick size, which might be expected to lead to less message volume, makes quote stuffing more difficult. Since the OER is a blunt instrument which may catch useful trading strategies as well, a higher minimum tick size might allow the OERs to be larger and still accomplish its role in reducing quote stuffing without inhibiting market making too much. If it was found that minimum tick size ought to be reduced for some securities because the spreads are artificially large for liquid stocks, then an OER may help to allow a smooth transition to lower tick sizes without an explosion of messages.

OERs and maker-taker pricing may be conflicting. If the OER is close to hitting its upper bound, it may also be possible for high frequency traders to 'create' executing trades if required. Such trading costs depend on the overall fees and rebates found in maker-taker pricing. For example, if a high frequency trader could establish that a tick inside the touch was empty, it could send both a buy and sell limit order at that price to the venue and get credit for two transactions. Alternatively, if a high frequency trader could predict accurately that it was at the head of the queue at the touch, it could trade with itself (at a high probability) by sending a matching marketable order at that precise time.

Internalisation and market maker obligations also interact. Unless compensated for market maker obligations, market makers need to make trading profits to outweigh their market making losses. But internalisation reduces the profits from active trading. Market maker obligations and minimum resting times clash in the sense that high frequency traders are required by market maker obligations to post limit orders with tight spreads while minimum resting times mean that other high frequency traders take advantage of those stale quotes. This may mean that high frequency traders snipe each other and that in volatile markets much of the trading would comprise high frequency traders trading with each other.

The following diagram illustrates some pairwise interactions between six regulatory measures (those currently being considered within MiFID II). This figure is only intended to be illustrative and not exhaustive. Also, it does not convey the detailed nature of the interactions, which will depend upon the specific way in which the various measures would be implemented. Some of these interactions are described in greater detail in the main text. It does, however, demonstrate that decisions concerning the implementation of such regulatory measures should not be taken in isolation, but that such interactions need to be carefully considered.

Figure 6.1: Illustrative example of key interactions between regulatory measures



To conclude, the main lesson of these dependencies is to underscore that whatever rules are implemented, they must be carefully calibrated against other parameters, such as the various tick sizes and the exact circuit breaking mechanisms in the primary exchanges.

6.13 Conclusions

This chapter has considered a variety of proposals to deal with the new world of computer-based trading in markets. In this chapter, we have summarised the views of studies directed towards understanding the impact of these proposed changes. We further summarise our position below.

The desirability of understanding algorithmic trading strategies and their impact on the market is laudable but achieving this through notification requirements, as, for example, currently envisioned in MiFID II, may not be feasible given the complexity of algorithms and their interactions in the market.

Circuit breakers have a role to play in high frequency markets, and they are found in virtually all major exchanges. Because of the interconnected nature of markets, however, there may be need for coordination across exchanges, and this provides a mandate for regulatory involvement at least in times of acute market stress. New types of circuit breakers triggered as problems loom rather than after they have emerged may be particularly effective in dealing with periodic illiquidity.

Tick size policy can have a large influence on transaction costs, market depth, and the willingness to provide liquidity. The current approach of allowing each European trading venue to choose its own minimum tick size has merits, but can result in unhealthy competition between venues and a race to the bottom. A uniform policy applied across all European[60] trading venues is unlikely to be optimal, but a coherent overall policy for minimum tick size that applies to subsets of trading venues may be desirable. This coordinated policy could be industry-based such the one agreed to by FESE members.

The current system of exchanges determining how to structure market maker obligations and pay for them seems to be working well for most markets. Extending those obligations more broadly across markets and to the market making function more generally is problematic.

The aim of a more stable limit order book is laudable, and minimum resting times seem a possible device to achieve that aim. Many of the independent academic authors have submitted studies which are very favourable to a slowing of the markets. Nevertheless, they are unanimously doubtful that minimum resting times would be a step in the right direction, in large part because such requirements favour aggressive traders over passive traders and so are likely to diminish liquidity provision.

An OER is a blunt measure that catches both abusive and beneficial strategies. It may not do too much harm if the upper limit is large enough not to hinder market making and intermediation, but to the extent that it is binding on those activities it may be detrimental to both spreads and liquidity. It is unlikely that a uniform OER across markets would be optimal because it depends upon the type of securities traded and the trader clientele in the market. If a ratio could be structured to target those quotes that are considered socially detrimental directly, then it might be a useful tool for combating market manipulation. The absence of research investigating costs and benefits, as well as the difficulty of actually setting this measure optimally, suggest caution in adopting this approach for the market.

The issue of maker-taker pricing is complex and related to other issues, such as order routing and best execution, which seems like a more promising way of constraining any negative effects of this practice.

The CLOB is of fundamental importance and everyone involved in equity trading has an interest in finding fruitful ways of improving the performance of the virtual central limit order book that currently operates in Europe. The debate on how best to improve shows opposing trends towards consolidation of trading venues and fragmentation. The tension between these two trends is likely to continue in the near future.

Internalisation of agency order flow is in principle a good thing for all parties concerned, especially where large orders are involved. However, the trend away from pre-trade transparency cannot be continued indefinitely without detrimental effects on the public limit order book and price discovery.

Call auctions are an alternative trading mechanism which would eliminate most of the advantage for speed in the electronic limit order book. They are widely used already in equity markets at open and close and following a trading halt. But no major market uses them exclusively, suggesting that they do not meet the needs of many traders. To impose call auctions as the only trading mechanism seems unrealistic as there are serious coordination issues related to hedging strategies that would make this undesirable.

Further to the regulatory measures outlined in this chapter, the Report needs to consider other factors; a discussion on financial transaction tax can be found in Box 6.1. Other considerations, such as trust in financial markets and standardisation are discussed in the next chapter.

60 This includes countries in the European Union (EU), the European Economic Area (EEA) and Switzerland.

Box 6.1: Financial transaction tax

On 28 September 2011 the European Commission put forward a detailed proposal for an EU-wide financial transaction tax[61]. The proposed tax covered a wide range of financial instruments and was much more comprehensive than the UK's stamp duty on shares. One of the stated objectives of the proposed tax was "to create appropriate disincentives for transactions that do not enhance the efficiency of financial markets thereby complementing regulatory measures aimed at avoiding future crises"[62]. The Commission's targets here included short-term trading, particularly automated and high frequency trading[63].

Since then, it became clear that the proposal would not achieve the unanimous support of EU Member States as required for it to be adopted across the EU[64]. A number of concerns were raised, including: the potential negative effect on GDP, in part resulting from an increase in the cost of capital and hence a decline in investment; the susceptibility of the tax to avoidance through relocation unless it was adopted at a global level; and the uncertainty as to the bearers of the economic incidence of the tax[65]. However, supporters of the proposal dismissed these concerns[66], and a number of Member States are currently considering implementing a financial transaction tax through the enhanced cooperation procedure. Subject to a number of conditions, this procedure allows a sub-group of Member States to introduce measures that only bind the participating states[67].

Within this Project we were not able to second guess the details of the tax that might eventually be adopted. These details matter since it will determine the impact of the tax.

Financial transaction taxes have been discussed, more generally, for many years. Apart from their revenue-raising potential, the perceived corrective function of these taxes is also often cited in their support. Proponents[68] thus argue that such taxes can produce positive effects on financial markets by increasing the cost and reducing the volume of short-term trading. Crucially, this argument is based on a view of short-term trading as being mostly speculative (often supported by or based on trading systems) and unrelated to market fundamentals. This form of trading is thus viewed as having a negative impact on financial markets by contributing to excessive liquidity, excessive price volatility and asset bubbles. Furthermore, it is argued that the increasing ratio of financial transactions to GDP suggests considerable socially unproductive financial activity and hence a waste of resources. Financial transaction taxes are also seen as a way of compensating for the fact that many financial services are exempt from VAT.

61 European Commission (2011a).

62 Ibid. p. 2. For an assessment of the proposal's objectives and whether a financial transaction tax is the instrument best suited to achieve them see Vella *et al.* (2011).

63 See, for example, European Commission (2011b) at pp. 27–28, 38 and 53. See also Semeta (2012). Commissioner Semeta here explained that "a second objective [of the FTT] is to discourage unwarranted and leveraged transactions such as high frequency trading which inflate market volumes in all segments. This should complement regulatory measures and expose financial market actors to price signals."

64 The opposition of some Member States, including Sweden and the UK, was evident from an early stage. See, for example, The Guardian (2011). The matter was put beyond doubt at a meeting of the European Union Economic and Financial Affairs Council (ECOFIN) on 22 June 2012.

65 For an outline of these concerns, as well as arguments made by supporters of the proposal, see the report by the House of Lords, European Union Committee (2012a). See also EIA21 (Annex D refers).

66 European Union Commission (2012) See also Griffith-Jones and Persaud, (2012).

67 Article 20 of the Treaty on the European Union and articles 326 to 334 of the Treaty on the Functioning of the European Union.

68 See, for example, Schulmeister *et al.* (2008).

In response it has been pointed out that: not all short-term trading is 'undesirable'; financial transaction taxes do not distinguish between long-term and short-term trading or between 'desirable' and 'undesirable' short-term trading; such taxes might not affect volatility or might even increase it because they reduce liquidity in markets; asset bubbles may also develop in the presence of high transaction costs, as documented by the recent real estate bubbles; and it is neither obvious what the ideal ratio of financial activity to GDP should be, nor is it clear whether this ratio should increase or decrease over time[69]. Finally, the VAT exemption of financial services reflects the difficulties of applying VAT to margin based financial services. Financial transaction taxes cannot solve this problem for a number of reasons. Most importantly, financial transactions taxes do not tax value added in the financial sector.

There have been a number of academic studies, both theoretical and empirical, on the effect of financial transaction taxes, and transaction costs more generally. These have been helpfully reviewed by Matheson (2011) and Hemmelgarn and Nicodeme (2010)[70]. Studies find that by increasing transaction costs, financial transaction taxes reduce trading volume. This generally reduces liquidity, which in turn can slow down price discovery[71]. The theoretical literature produces mixed results on the effect of financial transaction taxes on volatility, with a number of studies showing that market microstructure plays an important part[72]. The empirical literature on the effect of transaction costs on volatility also presents contrasting results, although studies with better data and estimation techniques seem to find more often a positive relationship between transaction costs and volatility. These studies mainly relate to short-term volatility[73]. Matheson points out that there is a lack of research on the relationship between transaction costs and long-term price volatility and price bubbles. However, transaction costs are unlikely to play a decisive role in determining market cycles[74]. Both theoretical and empirical studies generally find that transaction costs, including those by transaction taxes, tend to increase the cost of capital and lower asset prices[75]. There is also research using experimental methods to assess the effect of transaction taxes. In a recent study[76], the view that transactions are relocated to untaxed markets is confirmed, as is the finding that the impact on price volatility is in general ambiguous.

Overall, these results suggest that financial transactions taxes i) will give rise to significant relocation effects to tax havens and ii) cannot be expected to reduce price volatility in asset markets[77]. Yet there are important open questions, in particular the interaction between financial transaction taxes, other financial sector taxes, and regulation, and the effect of financial sector taxes on long term asset price movements and asset bubbles.

69 See for example, IMF (2010).

70 T. Matheson (2011); and Hemmelgarn & Nicodeme (2010).

71 See the studies reviewed in Matheson (2011) pp. 16–20. Note that some studies find that transaction taxes can reduce or increase liquidity depending on market microstructure: See for example, Subrahmanyam (1998).

72 See the studies reviewed in Hemmelgarn & Nicodeme (2010), pp. 34–36.

73 See Hemmelgarn & Nicodeme (2010) pp. 34–36 and the studies reviewed therein, and also the review in Matheson (2011) pp. 20–22.

74 Matheson (2011) p. 21.

75 See the studies reviewed in Matheson (2011) pp. 14–16.

76 Hanke *et al.* (2010).

77 See also IMF (2010) and the assessment in Anthony *et al.* (2012).

7 Computers and complexity

Key findings

Over coming decades, the increasing use of computers and information technology in financial systems is likely to make them more, rather than less complex. Such complexity will reinforce information asymmetries and cause principal/agent problems, which in turn will damage trust and make the financial systems sub-optimal. Constraining and reducing such complexity will be a key challenge for policy makers. Options for legislators and regulators include requirements for trading platforms to publish information using an accurate, high resolution, synchronised timestamp. Improved standardisation of connectivity to trading platforms could also be considered.

A particular issue is the vast increase in financial data, which are often not standardised, nor easily accessible to third parties (for example, regulators or academics), for analysis and research. The creation of a European financial data centre to collect, standardise and analyse such data should be considered.

7 Computers and complexity

7.1 Introduction

Computers play many roles in the financial field. The focus, in the preceding chapters, has primarily been on their speed which enables much faster (and usually more accurate and error free) transactions to be done by computers than by humans. This is the basis for the computer-based trading (CBT) on which much of this Report focuses. But information technology can provide far more than speed alone. Computers can hold, and allow access to, far more memory than could be stored in a combination of human brains and libraries. They can allow data not only to be accessed, but also to be manipulated and subjected to mathematical analysis to an extent that was previously unfeasible. New professional groups, such as quantitative analysts, or 'quants', in finance and various kinds of hedge funds have rapidly developed to exploit such possibilities.

Against such advantages from computerisation, there are also some drawbacks. In particular, the application of electronics and information technology makes finance not only more complex but also less personal, potentially endangering the framework of trust that has provided an essential basis to support financial dealing. These issues are discussed in Section 7.2. Steps that could be taken to reduce complexity and enhance efficiency, through greater standardisation of computer-based financial language and of financial data are discussed in Section 7.3.

7.2. Complexity, trust, and credit

There is a general and growing perception that financial arrangements, transactions and processes are not only complex, but have become much more so over recent decades. As Paul Volcker (2012) wrote[1]:

> Long before the 2008 financial crisis broke, many prominent, wise and experienced market observers used the forum provided by the now-renowned William Taylor Memorial Lecture – which I was recently privileged to give (Volcker 2011)[2], and around which this article is based – to express strong concerns about the implications of the greater complexity, the need to develop more sophisticated and effective approaches toward risk management and the difficult challenges for supervisors[3].

Complexity implies greater informational asymmetry between the more informed finance expert and the less informed client. Informational asymmetry in turn causes principal/agent problems where the informed principal is able to benefit at the expense of the uninformed client.

The experience of the financial crisis has underlined these concerns, through for example, the mis-selling of (synthetic) collateralised debt obligations (CDOs) to unwary buyers, and the mis-selling of certain kinds of borrowers' insurance. As a result of these and numerous similar examples, the financial industry has, to some considerable extent, lost the trust of the public. That, in itself, is serious.

Analysis on trust and regulation commissioned for this Project[4] includes the claim that computerisation and information technology reduces the need for personal interaction in financial services, and hence lessens the value of reputation and of longer term relationships. Several of the features of these changing services are, at least in part, a concomitant of computerisation. Today, the withdrawal of the personal touch, publicity about very substantial rewards to senior managers (especially after bank losses and failures), and reports of scandals, such as attempted London Interbank Offered Rate (LIBOR) fixing, have all served to erode trust.

1 Volcker (2012).

2 Volcker (2011).

3 For example, Cartellieri (1996); Crockett (1998); Fischer (2002).

4 DR30 (Annex D refers).

One reason why such loss of trust matters is that almost all financial deals involve some risks to each of the counterparties. If the uninformed client believes that the finance professional is always trying to abuse any trust, in order to maximise short-run profit, the client will not accept the deal so readily in the first place, and, if the deal is done, will be more ready to litigate on the basis of it not being appropriate if it goes wrong. In this environment there is a danger of a 'heads I win; tails I sue' culture developing, reinforced by a media and a judiciary who also doubt the morality of finance. The adverse effect that this can have on efficiency is, perhaps, best illustrated by the additional costs that litigation imposes on US medicine. The same trend could be developing in the finance industry.

For example, the best means of saving for pensions, according to finance theory, involves a mixture of equities and debt, with the mix depending on age. But equities are risky. So there can be a danger of financial advisers, in a world where trust has declined, putting clients into a portfolio consisting largely of 'riskless' government bonds and bank deposits, which has no reputational risk for the adviser, but has severe risks of inflation and a poor return for the client.

The basic financial instruments remain debt and equity. These are, in themselves, relatively easy to understand. Debt gives a fixed interest return; if this is not paid, the creditor gains certain rights of control over the assets of the debtor. The equity holder gets the residual return, after taxes, costs and interest. Given the uncertainty of estimating this residual return, the equity holder has underlying governance of the enterprise, in order to constrain inefficient or self interested management.

However, given limited liability, both equity holders and managers will find it advantageous to take on additional risk, and debt leverage, in pursuit of a higher return on equity (RoE)[5]. This factor was a major driver of the leverage cycle which has badly damaged the financial systems of the developed world over the past decade. In the absence of direct measures to restrict the incentives to take on 'excessive' leverage in boom periods, notably to remove the tax benefit of debt relative to equity and to reduce the built-in incentives of shareholders or managers to increase leverage in pursuit of RoE, the second best approach has been to advocate the introduction of hybrid instruments. These have some of the characteristics of debt, but can be like, or transform into, equity under certain (stressed) conditions.

Conditional Convertible debt instruments (CoCos) and bail-inable bonds for banks are a leading example. Shared Appreciation Mortgages (SAMs), sometimes also called 'Home Equity Fractional Interest', could be another. For government financing, GDP bonds, where the pay-out depends on the growth of the country's aggregate economy also has some similarities. Assuming no first-best reform of the tax system and of managerial/shareholder incentives, such hybrid instruments may become more common over the coming decade. They will also be considerably more complex, with the valuation of CoCos, for example, depending on many details such as the trigger and conversion terms. Moreover their introduction, could also complicate the calculation of the fundamental values of the remaining straight equities and bonds. This extra complexity, should it occur, will add to issues about what assets it should be appropriate for various kinds of savers to hold and for borrowers to issue.

A particular problem with SAMs is that when housing prices have risen, those that borrowed on such a basis will claim that they were misled, misinformed, mis-sold or unable to understand, so that they deserve to secure all the appreciation above the face value of the mortgage. The reputational and litigation risk of such an instrument is high. Yet housing finance is in real need of reform.

The investment decision that brought down Lehman Brothers was on the future direction of the housing market, and hence on the value of derivative instruments based on that market. Since World War II the major financial crises disrupting the UK economy have all been related to real estate bubbles and crashes, with their epicentre in commercial property. These events were the Fringe Bank crisis (1973–75), the 'boom and bust' (1988–1992), and the recent Great Moderation leading on to the Great Recession (2007 onwards). During the first part of these three periods rising property prices, both commercial and housing, were stimulated by laxer lending conditions, driven in most cases by competitive pressures from challenger, competitive banks aiming to transform themselves into major financial institutions. As the existing established oligopoly did not want to lose market share, the procyclical shifts in lending standards became exacerbated. When the bubbles eventually burst, the

5 See Admati *et al.* (2010 and 2012).

challenger banks were exposed (for example, Northern Rock in 2007), while the established banks were seriously weakened.

The subsequent fall in housing prices, combined with the difficulty of raising the subsequently increased down-payment, has typically led to a sharp decline in housing construction. The fall-off in house-building relative to potential housing demand then creates an excess, pent-up demand for housing, which then generates the next housing boom. On average there have been about 20 years between successive housing crises.

The procyclicality of housing and property finance needs to be dampened. While in principle this could be done via macro-prudential measures, the very considerable (political) popularity of the preceding housing booms makes it unlikely that such measures would be pressed with sufficient vigour. The unwillingness of the Financial Policy Committee (FPC) in the UK even to ask for powers to vary Loan to Value (LTV) or Loan to Income (LTI) ratios in the current state of public opinion and acceptance is a notable example. Alternatives may need to be sought in other forms of instrument, for example, the Danish covered bond mechanism[6].

For this, and other reasons, the financial system will likely become more, rather than less, complex in the future. Since such complexity reinforces information asymmetries, and causes principal/agent problems which damage trust and make the financial system sub-optimal, how can it be constrained and reduced?

This is a large subject. It includes education and disclosure, and the need for professional bodies to extract and to simplify the vast amounts of disclosed information that, assisted by information technology, is now routinely required and produced within the financial system, and which few have the time or inclination to read, as noted in the Kay Report (2012)[7]. One of the reasons for the popularity of Credit Rating Agencies (CRAs) was that they reduced the extensive information on credit-worthiness to a single statistic. Although they could have produced more qualifying information there was generally little demand from clients.

Information technology will allow the production of an even greater volume of data. While disclosure is generally beneficial to the users of finance, perhaps the greater problem from now on will be how to interpret the masses of data that are generated. How, for example, can the regulators identify evidence of market abuse from the vast amounts of data on orders, and transactions? Moreover, to detect market abuse, each national regulator will need access to international market data. Otherwise the market abuser can evade detection by transacting simultaneously in several separate national markets.

In the USA, the Office of Financial Research (OFR) has been commissioned by the Dodd-Frank Act to found a financial data centre to collect, standardise and analyse such data[8]. The increasing number of trading platforms, scattered amongst many locations, and the vast increase in the number of data points in each market, in large part a consequence of CBT, makes such an initiative highly desirable (see Chapter 5, *ibid*). There may be a case for a similar initiative to be introduced in Europe (see Chapter 8).

Confronted with this extra complexity in financial instruments, one proposal that is sometimes made is that (various categories of) agents should only be allowed to hold or issue instruments which have been approved by the authorities in advance. This contrasts with the more common position that innovation should be allowed to flourish, but with the authorities retaining the power to ban the uses of instruments where they consider evidence reveals undesirable effects. The former stance, however, not only restricts innovation, but also such official approval may well have unintended consequences.

6 In Denmark, mortgage banks completely eliminate market risk as the issued bonds perfectly match the loans granted. The linking of lending and funding is what has made the Danish mortgage system unique compared with other European mortgage systems. See https://www.nykredit.com/investorcom/ressourcer/dokumenter/pdf/NYKR_DanishCoveredBonds_WWW.pdf, and http://www.nykredit.com/investorcom/ressourcer/dokumenter/pdf/Danish_covered_bond_web.pdf Accessed: 17 September 2012.

7 Kay (2012).

8 See OFR (2012).

Many of the instruments that are sometimes accused of causing financial instability, such as credit default swaps (CDS), securitisations of various kinds, even sub-prime mortgages, were introduced to meet a need. Many, possibly all, of the instruments now condemned in some quarters as having played a part in the recent global financial crisis would at an earlier time have probably been given official approval. Officials have no more, probably less, skill in foreseeing how financial instruments will subsequently fare than CRAs or market agents.

Virtually any instrument that is designed to hedge can also be used to speculate. Since most assets are held in uncovered long portfolios, what really undermines assets when fear takes hold is usually the rush to hedge such uncovered long positions, rather than simple short speculation.

7.3 Standardisation

Financial complexity has increased, and will increase further. As described in the previous section it carries with itself many potential disadvantages and risks. There are, however, standardisation procedures that if adopted could help limit such complexity. Standards are important in both the financial services industry and the wider economy because they can help to engender trust, and to ensure levels of quality. They can allow businesses and consumers to realise economies of scale, and information to be more readily and transparently disclosed and gathered. Importantly they drive innovation by allowing interoperability to be achieved between systems. For example, standards govern the interoperability of the cellular networks for mobile telephone networks. Mobile telephones are charged using a standardised micro USB charger, the use of which is dictated by European and global standards. Standards contribute £2.5 billion in the UK's economy[9]. As such, the UK has an active position on standards and actively promotes their adoption and use. The need now is to do the same in finance as in these other industries.

World War II highlighted the need for international standards, sometimes with dramatic effect. The British Eighth Army fighting the panzers of Germany's Afrika Korps in the North Africa campaign was awaiting spare parts for their tanks from their American allies. However, they were found to be unusable because of different thread pitches on the nuts and bolts used in the parts. The British abandoned their tanks in the North African desert, as they were unrepairable[10, 11]. The American, British and Canadian governments subsequently joined forces to establish the United Nations Standards Coordinating Committee (UNSCC). After the war this initiative grew to include nations from neutral and enemy countries and in 1947 became the International Organisation for Standards (ISO) headquartered in Geneva[12].

Standards currently play a relatively limited role in financial services. There are some standards but those that do exist do not cover the broad spectrum of the whole infrastructure. The same levels of competition, innovation, transparency or the economies of scale do not appear in computer-based trading as in many other parts of the economy. One example is FIX which is a standard used for routing stock purchase order from investors to brokers, and reporting the execution of those stock purchases. Anecdotal evidence suggests that FIX helped establish electronic trading and central limit order book marketplaces. Where standards do exist in financial services they have largely been driven by private, rather than public benefits. Whilst standards exist for order routing, they came into existence primarily because large buy side institutions, paying substantial commissions to the broker dealers to whom they were routing orders, required them to implement the emerging standards, or risk seeing a reduction in their share of those commission payments.

9 Evidence base for the economic benefits of standardisation, http://www.bis.gov.uk/policies/innovation/standardisation/economic-benefits

10 Robb (1956) p. 526.

11 Surowiecki (2011).

12 Buthe & Mattli (2011).

For a standard to be widely adopted it needs to satisfy two criteria; first, it needs to be open in the sense that it can be readily and easily adopted by all participants in the marketplace. Second, market participants need to believe that the standard is credible (credibility here means that people believe that the standard will be widely adopted). The credibility of a standard may be undermined by external factors such as the Intellectual Property Rights (IPR) framework within which standards are delivered. DR31[13] considers standards in more detail, reviews their role in the economy, their existence in financial services and suggests how they may play a larger role in delivering transparency and trust, and reducing risk of financial instability.

Standards are needed in areas of the market where they will help deliver public benefits such as greater competition, transparency, innovation and financial stability, yet often they only arise in those areas where their adoption will defend the economic interests of participants in the market place. Committees, made up largely of banks and vendors, currently govern standards and there is little representation of other stakeholders, the investors, regulators or wider society. A mandate for standardisation initiatives which are deemed to serve the wider public interest rather than the narrower private interest could help secure the credibility to enable widespread adoption.

More effective governance should ensure that standards address the interests of all participants in financial markets. Identification of areas with either high risk to financial stability or high cost implications for end investors is urgently required. Particular attention is needed where information asymmetries are causing market failure and systems weakness.

Specific requirements could include:

- All trading platforms to share and publish information to an accurate, high resolution, synchronised timestamp. This could make surveillance easier across global markets and simplify comparisons of trading which takes place across multiple fragmented venues.
- The exploration of the potential for common gateway technology standards to ease platform connectivity and improve the environment for collecting data required for regulatory purposes. Clearly, common gateway technology standards for trading platforms could enable regulators and customers to connect to multiple markets more easily, making more effective market surveillance a possibility. They could also create a more competitive market place in surveillance solutions.
- A review of engineering processes for electronic secondary markets. Areas that create artificially high costs within the industry and areas that increase the risk of financial instability in transaction processing need identifying. This information could inform regulation and standardisation.
- The creation of a standard process for making order book data available to academic researchers. This step must address the concerns of the industry on privacy but also facilitate the conduct of academic research. This will contribute to the creation of an evidence base to help assess the need for further action.

The initiatives of the G20 in mandating the introduction of a legal entity identifier (LEI) and universal product identifier are good first steps[14].

7.4 Conclusions

The financial sector over expanded in the early years of the current century, with excessive leverage superimposed on insufficient equity. This leverage was partly enabled by complex innovations such as securitisation, which was in turn enabled by information technology. If the incentives for excessive leverage cannot be directly addressed, more hybrid instruments, which combine both debt and equity characteristics, may usefully be introduced. This may be so especially in the field of housing finance, where shortcomings in the latest crisis have not yet been adequately addressed.

This additional complexity has reinforced a breakdown of trust which in turn has damaged the operations of finance. A corrective step that could, and should, be taken, is to simplify (electronic) financial systems by the application of greater standardisation, particularly in the form of an accurate,

13 DR31 (Annex D refers).

14 Financial Stability Board (2012).

high resolution synchronised timestamp. Moreover, CBT, operating on many trading platforms, has led to a vast expansion of data, which are often not standardised, nor easily accessible to third parties (such as, regulators and academics) for analysis and research. The relevant authorities should consider following the US example and establish a European Financial Data Centre to collect, standardise and analyse such data.

Chapter 8 draws together the conclusions across the Report to provide policy makers with the key findings and future options.

8 Conclusions and future options

Key findings

Analysis of the available evidence has shown that computer-based trading (CBT) has several beneficial effects on markets, notably relating to liquidity, transaction costs and the efficiency of market prices. Against the background of ever greater competition between markets, it is highly desirable that any new policies or market regulations preserve these benefits. However, this Project has also highlighted legitimate concerns that merit the close attention of policy makers, particularly relating to the possibility of instabilities occurring in certain circumstances, and also periodic illiquidity. The following suggests priorities for action:

A: Limiting possible future market disturbances:

- European authorities, working together, and with financial practitioners and academics, should assess (using evidence-based analysis) and introduce mechanisms for managing and modifying the potential adverse side-effects of CBT.
- Regulatory constraints involving CBT in particular need to be introduced in a coordinated manner across all markets where there are strong linkages.
- Regulatory measures for market control must also be undertaken in a systematic global fashion to achieve in full the objectives they are directed at. A joint initiative from a European Office of Financial Research and the US Office of Financial Research (OFR), with the involvement of other international markets, could be one option for delivering such global coordination.
- Legislators and regulators need to encourage good practice and behaviour in the finance and software engineering industries. This clearly involves the need to discourage behaviour in which increasingly risky situations are regarded as acceptable, particularly when failure does not appear as an immediate result.
- Standards should play a larger role. Legislators and regulators should consider implementing accurate, high resolution, synchronised timestamps because these could act as a key enabling tool for analysis of financial markets. Clearly it could be useful to determine the extent to which common gateway technology standards could enable regulators and customers to connect to multiple markets more easily, making more effective market surveillance a possibility.
- In the longer term, there is a strong case for lessons to be learnt from other safety critical industries, and to use these to inform the effective management of systemic risk in financial systems.

B: Making surveillance of financial markets easier:

- The development of software for automated forensic analysis of adverse/extreme market events would provide valuable assistance for regulators engaged in surveillance of markets. This would help to address the increasing difficulty that people have in investigating events.

C: Improving understanding of the effects of CBT in both the shorter and longer term:

- Unlocking the power of the research community has the potential to play a vital role in addressing the considerable challenge of developing better evidence-based regulation relating to CBT risks and benefits and also to market abuse. Suggested priorities include:
  - developing an 'operational process map': this would detail the processes, systems and interchanges between market participants through the trade life cycle, and so help to identify areas of high systemic risk and broken or failing processes;
  - making timely and detailed data across financial markets easily available to academics, but recognising the possible confidentiality of such data.
- The above measures need to be undertaken on an integrated and coordinated international basis to realise the greatest added value and efficiency. One possible proposal would be to establish a European Financial Data Centre.

8 Conclusions and future options

8.1 Introduction

Well functioning financial markets are vital for the growth of economies, and the prosperity and well-being of individuals. They can even affect the security of entire countries. But markets are evolving rapidly in a difficult environment, characterised by converging and interacting macro- and microeconomic forces. The development and application of new technology is arguably causing the most rapid changes in financial markets. In particular, high frequency trading (HFT) and algorithmic trading (AT) in financial markets have attracted controversy. While HFT and AT have many proponents, some believe they may be playing an increasing role in driving instabilities in particular markets or imposing a drag on market efficiency. Others have suggested that HFT and AT may encourage market abuse.

What is clear is that new capabilities will include the assimilation and integration of vast quantities of data, and more sophisticated techniques for analysing news. There will be increasing availability of substantially cheaper computing power, particularly through cloud computing. The extent to which different markets embrace new technology will critically affect their competitiveness and therefore their position globally: those who embrace this technology will benefit from faster and more intelligent trading systems. Emerging economies may come to threaten the long-established historical dominance of major European and US cities.

This Report has sought to determine how computer-based trading (CBT) in financial markets could evolve by developing a robust understanding of its effects. Looking forward ten years, it has examined the evidence to identify the risks and opportunities that this technology could present, notably in terms of financial stability[1] but also in terms of other market outcomes such as volatility, liquidity, price efficiency and price discovery, and the potential for market abuse.

In this final chapter, the conclusions and the options for policy makers are presented in three sections. They comprise:

- an overall assessment of whether HFT/CBT has imposed costs or benefits. The challenges of understanding the behaviour of CBT financial markets are then considered;
- conclusions on policies that are being considered by policy makers with the goals of improving market efficiency and reducing the risks associated with financial instability;
- consideration of how the connectivity and interactions of market participants can be mapped and monitored so that the behaviour of CBT in financial markets can be better understood, and regulation better informed.

8.2 HFT: the pursuit of ever greater speed in financial transactions

In sport, the ideal is to benefit from 'taking part'; in practice, the aim is, almost always, to come first and win (gold). In financial markets, the ideal is to discover the 'fundamental' price of assets; in practice, the aim is, almost always, to react first to any 'news' that might drive financial prices and thereby win gold (of a more pecuniary nature). Whether there is any social value in running some distance .001 of a second faster, or reducing the latency of an electronic response to news by .00001 of a second, is a moot point. But what is clear is that both efforts are driven on by basic and natural human instincts. Trying to prevent the attempt to learn, to assess, to respond and to react to news faster than the next agent would only drive CBT, trading volumes and liquidity to some other venue, where such assets can be freely traded. And that market would come to dominate price setting.

1 A list of definitions used in this chapter can be found in Annex C.

While the effect of CBT on market quality is controversial, the evidence suggests that CBT has generally contributed to improvements. In particular, over the past decade, despite some very adverse macroeconomic circumstances:

- liquidity has improved;
- transaction costs have fallen for both retail and institutional traders;
- market prices have become more efficient, consistent with the hypothesis that CBT links markets and thereby facilitates price discovery.

In the European Union (EU) the Markets in Financial Instruments Directive (MiFID) led, as intended, to the establishment of a wide variety of competitive trading platforms, many of these having innovative features. Now that the same asset can be traded on many different platforms, there is a need for arbitrage to interconnect them and HFT provides this. Indeed the growth of HFT was, in some part, an unintended consequence of MiFID. The bulk of the evidence reviewed in Chapter 3 suggests that, in normal circumstances, HFT enhances the quality of the operation of markets, though the returns to HFT may come in some significant part from buy-side investors, who are concerned that HFT may have the effect of moving prices against them:

- Greater transparency should allow testing of this conclusion. The introduction of a synchronised timestamp for electronic trading would help achieve this (see Section 8.4 for further discussion).

It is not so much concerns from these investors, but rather fears that, under stressed conditions, the liquidity provided by HFT may evaporate and lead to greater volatility and more market crashes, which provides the main driving force for new regulatory measures to manage and to modify CBT. **It is clear that there are various mechanisms whereby self-reinforcing feedback loops can develop, leading to market crashes, and that HFT and CBT could be a factor in some of these mechanisms (see Chapter 4**).

However, several caveats apply:

- **Crashes occurred prior to HFT** (19 October 1987 being a prime example). The idea that market makers would, and could, previously always respond to dampen 'fast' markets is not supported by analysis of past events.
- **HFT is a recent phenomenon and market crashes remain rare events**. There is contention about the cause of the Flash Crash on 6 May 2010. There is as yet insufficient evidence to determine what role HFT played either in the Flash Crash, or other mini-crashes that have occurred since HFT became established.
- **A key element of HFT, and CBT more broadly, is that they link markets together.** If a (regulatory) constraint is put on one market while another market is free of that constraint, then, especially in stressed conditions, electronic orders and trading will migrate to the unconstrained market. If that market is thinner, prices will move even more, and those price changes will react back to the constrained market.

8.3 Managing and modifying CBT

There are a variety of policies being considered by policy makers with the goals of improving market efficiency and reducing the risks associated with financial instability. This is the main subject of Chapter 6, where all the main proposals have been carefully assessed with a particular focus on their economic costs and benefits, implementation issues and empirical evidence on effectiveness.

Overall, there is general support for circuit breakers, particularly for those designed to limit periodic illiquidity induced by temporary imbalances in limit order books. Different markets may find different circuit breaker policies optimal, but in times of overall market stress there is a need for coordination of circuit breakers across markets, and this could be a mandate for regulatory involvement. Further investigation is needed to establish how this could best be achieved in the prevailing market structure.

There is also support for a coherent tick size policy across similar markets. Given the diversity of trading markets in Europe, a uniform policy is unlikely to be optimal, but a coordinated policy across competing venues may limit excessive competition and incentivise limit order provision.

There is less support for policies imposing market maker obligations and minimum resting times on orders. The former issue runs into complications arising from the nature of high frequency market making across markets, which differs from traditional market making within markets. Requirements to post two-sided quotes may restrict, rather than improve, liquidity provision. Similarly, minimum resting times, while conceptually attractive, can impinge upon hedging strategies that operate by placing orders across markets and expose liquidity providers to increased 'pick-off risk' if they are unable to cancel stale orders.

Proposed measures to require notification of algorithms or minimum order-to-execution ratios are also not supported. The notification policy proposal is too vague, and its implementation, even if feasible, would require excessive costs for both firms and regulators. It is also doubtful that it would substantially reduce the risk of market instability due to errant algorithmic behaviour. An order-to-execution ratio is a blunt instrument to reduce excessive message traffic and cancellation rates. While it could potentially reduce undesirable manipulative trading strategies, it may also curtail beneficial strategies. There is not sufficient evidence at this point to ascertain these effects, and so caution is warranted. Explicit fees charged by exchanges on excessive messaging and greater regulatory surveillance geared to detect manipulative trading practices may be more desirable approaches to deal with these problems.

Other policies remain problematic. Overall, the limited evidence suggests that maker-taker pricing improves depth and trading volume without negatively effecting spreads. However, the issue of maker-taker pricing is complex and related to other issues, such as order routing and best execution, which seems like a more promising way of constraining any negative effects of this practice. The central limit order book is of fundamental importance and everyone involved in equity trading has an interest in changes that improve the performance of the virtual central limit order book currently operating in Europe. The debate on how best to do so continues, in part because the opposing trends towards consolidation of trading venues and fragmentation is likely to remain unresolved in the near future.

Internalisation of agency order flow, in principle, benefits all parties involved, especially where large orders are involved. However, the trend away from pre-trade transparency cannot be continued indefinitely without detrimental effects on the public limit order book and price discovery. Call auctions are widely used already in equity markets at open and close and following a trading halt. But no major market uses call auctions exclusively to trade securities. To impose them as the only trading mechanism seems unrealistic as there are serious coordination issues related to hedging strategies that make this undesirable.

It should be recognised that some of the above individual policy options interact with each other in important ways. For example, the presence or absence of circuit breakers affects most other measures, as does minimum tick sizes. Decisions on individual policies should not therefore be taken in isolation, but should take account of such important interactions[2].

Coordination of regulatory measures between markets is important and needs to take place at two levels:

- **Regulatory constraints involving CBT in particular need to be introduced in a coordinated manner across all markets where there are strong linkages.**
- **Regulatory measures for market control must also be undertaken in a systematic global fashion to achieve in full the objectives they are directed at.** A joint initiative from a European Office of Financial Research and the US Office of Financial Research (OFR), with the involvement of other international markets, could be one option for delivering such global coordination.

2 See Chapter 6, Section 6.12 and also the supporting evidence papers which were commissioned (Annex D refers).

8.4 Mapping and monitoring CBT

A major challenge for understanding the behaviour of today's heavily computer-based financial markets is that there is no reference 'system map' of the network of connectivity and interactions among the various significant entities within the financial markets. Consequently, there can be severe problems in gaining access to appropriately accurate data concerning key relationships and sequences of events in such connected markets. Volumes of market data have grown significantly in recent years, and a system map would need to be constantly updated to reflect the dynamic development of the financial networks, further increasing issues of data bandwidth.

Dealing with such 'big data' presents significant new challenges. These difficulties are faced not only by regulatory authorities responsible for governing the markets, but also by academics attempting to advance our understanding of current and future market systems. **Market data needs to be collected, made available for study, analysed and acted upon in a timely fashion.**

There is also a need to have a reference 'operational process map' of the processes that act over the networks of the financial system. This would set out the processes, systems and interchanges between market participants throughout the trade life cycle. This map should aim to identify the areas of high systemic risk and those areas with broken or failing processes. Policy makers' attention should be directed at the areas highlighted to determine if a policy response is appropriate, to avoid systemic failure.

Standards currently play a relatively limited role in financial services. Certain standards are in place but those that do exist do not cover the broad spectrum of the whole infrastructure. Particular attention to standardisation is needed for areas where information asymmetries are causing market failure and systems weakness (see Chapter 7). In particular, **the development of accurate, high resolution, synchronised, timestamps would make the consolidation of information from multiple sources easier and also promote greater transparency.** The former, in particular, is a priority to enable data from diverse sources to be easily consolidated into a view that can be used as an audit trail in any location. The use of synchronised clocks for time-stamping market event data across key European trading venues would need to be used, so that the event-streams from European trading venues can be brought together. This would then allow the creation of high quality European public-domain datasets to inform policy analysis and investment strategy. Clearly it could be useful to determine the extent to which common gateway technology standards would ease platform connectivity and improve the environment for collecting data for regulatory purposes.

Attempting to map and monitor CBT piecemeal, country by country, or market by market, would be wasteful and inefficient. **There would be value in establishing a European Financial Data Centre (EFDC).** One purpose of this would be to ensure that, not only is sufficient data collected and kept but, just as important, that data emanating from differing European countries are compatible and consistent.

Besides improving the collection and consistency of financial data, an EFDC would be the natural source whereby academics could access and utilise the data. That should be a two-way process since academic needs and proposals should help to inform the construction and continuing evolution of the financial databases.

8.4.1 Monitoring CBT: analysing the data

Further research is required to develop appropriate statistical methodology for analysing high frequency, multi-venue order book data for multiple instruments over multiple trading periods. These methods could then be used to evaluate different policy options. They would allow accurate tracking of any shortfalls in implementation, and the evolution of market quality, facilitate the monitoring of transaction cost analysis, and would also increase the opportunities for *ex-post* detection of market abuse.

The development of automated real-time monitoring of markets is likely to be very difficult to implement and maintain effectively. **Instead, it is likely to be more cost effective to develop software for automated forensic analysis of adverse/extreme market events, after they have occurred, particularly if all markets share a common timestamp, and connectivity has been simplified.** Establishing timely causal explanations for such events can be very challenging for experts, even if all relevant market data have been captured and stored for subsequent analysis. If (as may be expected) some data are missing or corrupted, the challenges increase significantly.

One possible exercise for an EOFR would be to build a simulator where a continuous-time limit order book is modelled and tested with a variety of algorithms to study the systemic properties of the nonlinear interactions between the algorithms. This would offer an opportunity for readily replicable empirical studies. Currently, not much is known about the interactions among algorithmic CBT systems in normal and in stressed conditions. Since the stability of markets and the confidence in the markets depends on the outcome of the markets' continuous double auction, analyses of the impacts of various algorithms or of various policy actions could be informed by results from simulation models. Simulators could also be constructed to help study not only interactions among CBT algorithms, but also the dynamics of markets populated both by CBT systems and by human traders (see DR25 and DR27).

In essence, simulation tools and techniques could enable central regulatory authorities to judge the stability of particular financial markets, given knowledge of the structure of those markets (i.e. the network of interacting financial institutions: who interacts with whom) and the processes operating on those networks (i.e. who is doing what, and why). A family of simulation models of various types (for example, agent-based simulations, and/or more analytical quantitative risk assessment techniques) would need to be constructed. For researchers the challenge is how best to model and predict the dynamics of existing markets, and how best to evaluate different management approaches and different potential modifications (for example, via running 'what-if' simulation models), before they are enacted in the real world. The example of the OFR in the USA which comes under the aegis of the Financial Stability Oversight Council (FSOC), and its data centre provides one model. The appropriate oversight body for an EOFR would be a matter for discussion.

Finally, 'normalisation of deviance' was introduced in Chapter 2 and discussed further in Chapter 4 as a potential destabilising factor in financial markets which are heavily dependent on technology that is 'risky' in the technical sense that the limits of safe operating behaviour and the vulnerabilities of the technology are not fully known in advance. Dealing with normalisation of deviance is problematic for policy makers because it is a social issue, rooted in the culture of organisations and in failures of communication within and between organisations. Normalisation of deviance occurs in situations where circumstances force people to learn the technology's limits by experience, and where organisational culture does not place a sufficiently high premium on the avoidance of adverse events and accidents. Altering social behaviour and organisational culture via regulatory means can be difficult.

Legislators and regulators should consider how best to encourage desired accident-avoiding practices and behaviours, and to discourage the normalisation of deviance. Safety-critical engineering practices have been developed in aerospace, in nuclear power engineering and similar fields, in part because of the reputational damage and legal repercussions which can follow from serious accidents. Studies of high-reliability organisations (HROs) were discussed briefly in Chapter 2 in the context of working practices that avoid normalisation of deviance. These practices have been developed and refined in HROs in part because of the culture and ethics of the organisations, but also because of the fear of punitive consequences where the failure of a process can subsequently be shown to be professionally negligent, or even criminal. Nevertheless, because normalisation of deviance occurs in situations where people are learning the limits of a system via experience, providing incentives may not be sufficient.

Internal working practices common in HROs and other areas of safety-critical systems engineering are not yet widespread in the engineering teams responsible for CBT systems in financial institutions. Similarly, financial regulatory authorities seem likely to benefit from better understanding of regulating safe systems in, for example, air transport or medical engineering. **What is needed in the finance industry is a greater awareness of the systemic risks introduced by normalisation of deviance, and more dialogue with, and opportunities to learn from, other industries where safety-critical and high-integrity engineering practices have been developed over decades.** Without this, there is a possibility that systemic risks are heightened, because there will be little or no way to predict the systemic stability of financial markets other than by observing their failures, however catastrophic those failures turn out to be for individual investors or financial institutions.

Finally, computerisation has implications for the conduct of financial markets beyond its effect on the speed of transactions. In Chapter 7 some of these wider implications are, briefly, considered, particularly the potential role of computerisation in eroding trust in financial markets. While this may be the case, standards apart, it is hard to think of policies to restore trust beyond those already put forward in the Kay Report[3].

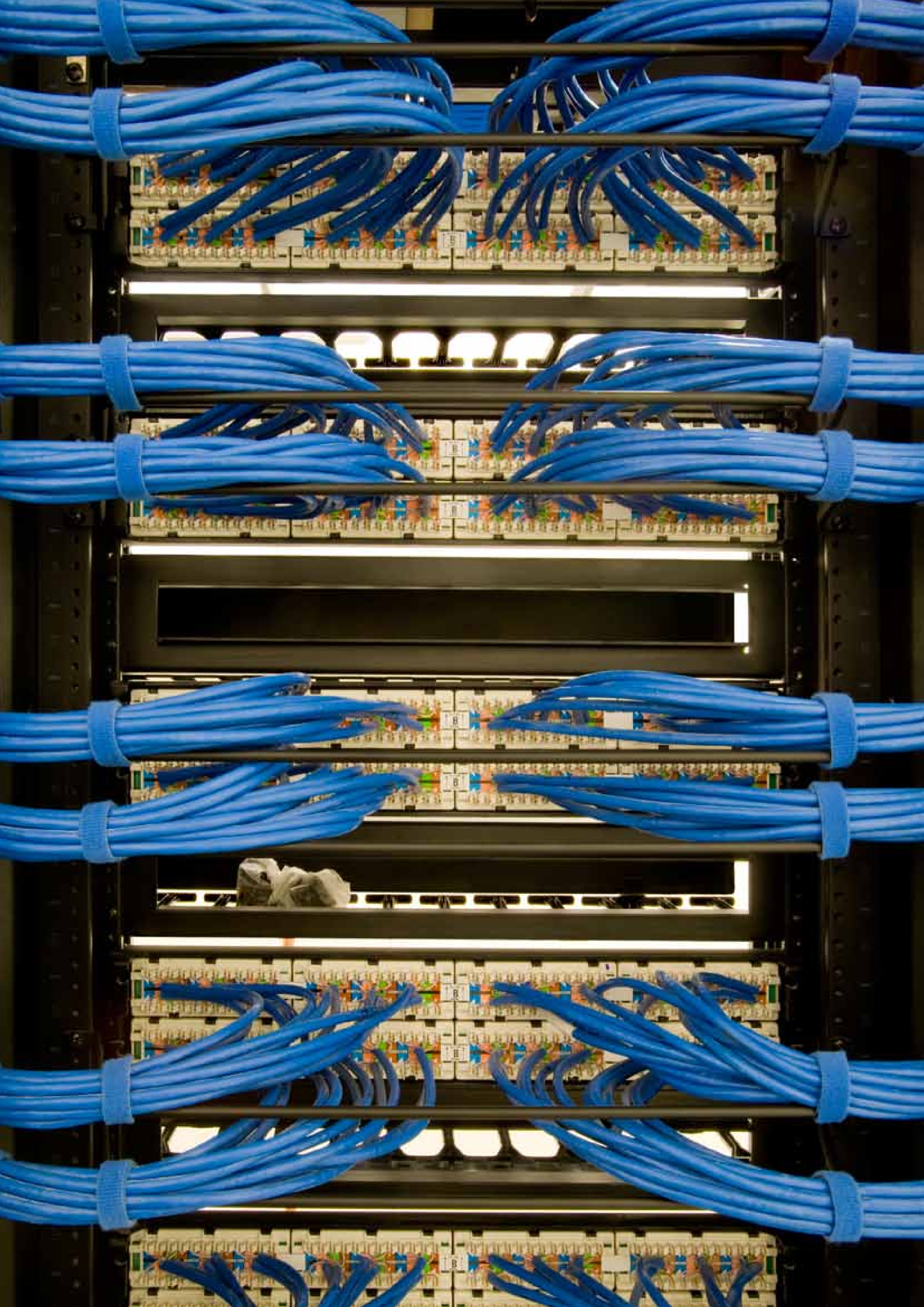
8.5 Conclusions

This report has presented evidence that CBT can improve the quality of markets, fostering greater liquidity, narrowing spreads, and increasing efficiency. Yet these benefits may come with associated costs: the rates at which current systems can interact autonomously with each other raises the risk that rare but extreme adverse events can be initiated and then proceed at speeds very much faster than humans can comfortably cope with, generating volumes of data that can require weeks of computer-assisted analysis by teams of skilled analysts before they are understood. Although adverse events may happen only very rarely, there is a clear danger that very serious situations can develop at extreme speed.

This Report has suggested options for limiting possible future market disturbances, making surveillance of financial markets easier, and for improving the scientific understanding of how CBT systems act and interact in both the shorter and the longer term. All of the options have been considered in the context of a substantial body of evidence drawn from peer reviewed research studies commissioned by the Project from a large international group of experts. The options presented here are directed not only to policy makers and regulators but also to market practitioners.

The world's financial markets are engines of economic growth, enabling corporations to raise funds and offering investors the opportunity to achieve their preferred balance of expected risks and rewards. It is manifestly important that they remain fair and orderly. Deciding how best to ensure this, in light of the huge growth in both the uptake and complexity of CBT that has occurred in the last decade, and which can be expected to continue in the next, requires careful thought and discussion between all interested parties.

3 Kay (2012).



Annex A Acknowledgements[1]

The Government Office for Science would like to express its thanks to the following individuals who were involved in the Project's advisory bodies, undertook the detailed technical work and who provided expert advice in the Project's many workshops and studies. Foresight would also like to thank the individuals who peer-reviewed individual papers, and the many others not listed here who contributed views and advice in other ways.

HIGH-LEVEL STAKEHOLDER GROUP

Dr	Greg	Clark MP (Chair)	Financial Secretary, Her Majesty's Treasury
	Mark	Hoban MP	Former Financial Secretary, Her Majesty's Treasury and Former Chair
	Ronald	Arculli OBE	Chairman, Hong Kong Exchanges and Clearing
Dr	John	Bates	Executive Vice President and Chief Technology Officer, Progress Software
Prof. Sir	John	Beddington CMG FRS	Government Chief Scientific Advisor, Government Office for Science
Dr	Adrian	Blundell-Wignall	Deputy Director (Financial and Enterprise Affairs), OECD
	Andrew	Bowley	Managing Director, Head of Electronic Trading Product Management, Nomura
	Dominique	Cerutti	President, Deputy CEO and Head of Global Technology, NYSE Euronext
	Stijn	Claessens	Assistant Director, Research Department, IMF
	Scott	Cowling	Global Head of Scientific Equity Trading, Blackrock
Dr	Claire	Craig CBE	Deputy Head of the Government Office for Science
	Kerim	Derhalli	Managing Director and Global Head of Equity Trading and Head of Equities EMEA, Deutsche Bank
	Richard	Gorelick	CEO, RGM Advisors LLC
	Philippe	Guillot	Executive Director of the Markets Directorate, Autorité des Marchés Financiers
	David	Harding	CEO, Winton Capital
	Christopher	Jackson	Director (Execution Sales), Merrill Lynch
	Christopher	Marsh	Managing Director (Advanced Execution Services), Credit Suisse
	John	McCarthy	General Counsel, GETCO
	Robert	McGrath	Global Head of Trading, Schroders
	Nicholas	Nielsen	Global Head of Trading, Marshall Wace
	Mark	Northwood	Global Head of Equity Trading, Fidelity
	Stephen	O'Connor	Chairman, International Swaps & Derivatives Association
	Scott	O'Malia	Commissioner, Commodity Futures Trading Commission
	Olivier	Osty	Deputy Head of Global Equities & Commodity Derivatives, BNP Paribas
	Jon	Robson	President of Enterprise, Thomson Reuters
	Xavier	Rolet	CEO, London Stock Exchange
	Thomas	Secunda	Founding Partner, Bloomberg
Dr	Kay	Swinburne MEP	European Parliament
	Ali	Toutounchi	Managing Director (Index Funds), Legal & General Investment Management
	Danny	Truell	Chief Investment Officer, Wellcome Trust
	Paul	Tucker	Deputy Governor, Bank of England
	Laurent	Useldinger	CEO, ULLink
	Martin	Wheatley	Managing Director, Consumer and Markets Business Unit, FSA

1 Individuals are only listed once. Where an individual contributed in multiple ways (for example, producing an evidence paper and attending a workshop) they are only listed for their first recorded contribution.

LEAD EXPERT GROUP[2]

Dame	Clara	Furse (Chair)	Non-executive Director, Legal & General plc, Amadeus IT Holding SA, Nomura Holdings Inc., Chairman, Nomura Bank International, Non-executive Director, Department for Work and Pensions and Senior Adviser, Chatham House.
Professor	Philip	Bond	Visiting Professor of Engineering, Mathematics and Computer Science at the University of Bristol and Visiting Fellow at the Oxford Centre for Industrial and Applied Mathematics.
Professor	Dave	Cliff	Professor of Computer Science at the University of Bristol.
Professor	Charles	Goodhart CBE, FBA	Professor (Emeritus) of Banking and Finance at the London School of Economics.
	Kevin	Houstoun	Chairman of Rapid Addition and co-Chair, Global Technical Committee, FIX Protocol Limited.
Professor	Oliver	Linton FBA	Chair of Political Economy at the University of Cambridge.
Dr	Jean-Pierre	Zigrand	Reader in Finance at the London School of Economics.

Foresight would like to thank Dr Sylvain Friederich, University of Bristol, Professor Maureen O'Hara, Cornell University and Professor Richard Payne, Cass Business School, City University, London for their involvement in drafting parts of this Report.

Foresight would also like to thank Andy Haldane, Executive Director for Financial Stability at the Bank of England, for his contribution in the early stages of the Project.

AUTHORS AND CONTRIBUTORS TO THE EVIDENCE BASE – DRIVER REVIEWS

Professor	Khurshid	Ahmad	Trinity College, Dublin
Professor	Michael	Aitken	University of New South Wales
Dr	Ali	Akyol	The University of Melbourne
Professor	Peter	Allen	Cranfield University
Professor	James	Angel	Georgetown University
	Angelo	Aspris	The University of Sydney Business School
Professor	Robert	Battalio	Finance and Presidential Faculty Fellow at the University of Notre Dame
Professor	Alessandro	Beber	City University, London
Professor	Itzhak	Ben-David	The Ohio State University
Dr	Daniel	Beunza	The London School of Economics
Professor	Bruno	Biais	The Toulouse School of Economics
Professor	Robin	Bloomfield	City University, London
Professor	Ekkehart	Boehmer	Ecole des Hautes Etudes Commerciales du Nord
Dr	Jonathan	Brogaard	University of Washington
Professor	Seth	Bullock	School of Electronics and Computer Science, University of Southampton

2 The following acknowledges associations of the Lead Expert Group members:

- **Dame Clara Furse** is a non-executive Director of Nomura Holdings, Legal & General Group, Amadeus IT Holdings and the Department for Work and Pensions. Dame Clara is also a Senior Advisor to Chatham House.
- **Professor Phillip Bond** has not declared any association that has bearing on their involvement with the Project.
- **Professor Dave Cliff** has not declared any association that has bearing on their involvement with the Project.
- **Professor Charles Goodhart** is a part-time Senior Economist Consultant to an investment bank in London and an occasional economic adviser to a couple of financial institutions in London, all of which may, and some of which do, use CBT techniques.
- **Kevin Houstoun** is a part-time Chairman of Rapid Addition Limited, http://www.rapidaddition.com/, a leading vendor of messaging components to the trading industry. Many of Rapid Additions clients use CBT techniques. Kevin also consults to investment banks, stock exchanges, vendors and the buy side on messaging and strategy; aspects of this work include CBT. Kevin has an minority investment in Thematic Networks, http://www.thematic.net/, the leading supplier of bespoke enterprise social network platforms. Thematic networks has investments in many trade publications including HFT Review, http://www.hftreview.com/, and Money Science, http://www.moneyscience.com/. Kevin was a partner in Fourth Creek, a FX high frequency trading firm, prior to its sale to ICAP in 2008. Kevin is actively involved in standardisation within the trading space and current co-Chairs the FIX Protocol Limited Global Technical Committee, he has held this position since 2004.
- **Professor Oliver Linton** has not declared any association that has bearing on their involvement with the Project.
- **Dr Jean-Pierre Zigrand** has not declared any association that has bearing on their involvement with the Project.

Professor	Muffy	Calder OBE, FRSE	University of Glasgow
Dr	John	Cartlidge	University of Bristol
Professor	Carole	Comerton-Forde	Australian National University
Professor	Gregory	Connor	National University of Ireland
Professor	Rama	Cont	Columbia University
Professor	Douglas	Cumming	York University
Dr	Jon	Danielsson	London School of Economics
	Marco	de Luca	Budget Planning and Control Analyst at Alenia North America
Professor	Doyne	Farmer	Santa Fe Institute
Professor	Karl	Felixson	Hanken School of Economics
	Sean	Foley	Capital Markets Cooperative Research Centre
	Kingsley	Fong	Australian School of Business
Professor	Thierry	Foucault	Ecole des Hautes Études Commerciales de Paris
Dr	Sylvain	Friederich	University of Bristol
Professor	Alex	Frino	University of Sydney
	Gérard	Gennotte	Ecole d'Economie de Paris
Professor	Michael	Goldstein	Babson College
Professor	Carole	Gresse	Université Paris-Dauphine
Professor	Antonio	Guarino	University College, London
Dr	Greg	Gyurko	University of Oxford
Professor	Frederick	H.deB. Harris	Wake Forest University
Professor	Terrence	Hendershott	University of California, Berkeley
Professor	David	Hutchison	Lancaster University
Professor	Giulia	Iori	City University, London
	Kevin	James	
Professor	Nicholas	Jennings	Southampton University
Professor	Neil	Johnson	Miami University
Professor	Charles	Jones	Columbia University
Dr	Guray	Kucukkocaoglu	Baskent University
	Torben	Latza	University of Bristol
Professor	Hayne	Leland	University of California, Berkeley
Professor	Marc	Lenglet	Head of Department (Management and Strategy) at European Business School
Professor	Andrew	Lepone	University of Sydney
Professor	Terry	Lyons	University of Oxford
Professor	Katya	Malinova	University of Toronto
Professor	Konstantinos	Manolarakis	Visiting researcher at Imperial College, London; MANOLOTECH LTD
Professor	Peter	McBurney	King's College, London
Professor	John	McDermid	Leader of the High Integrity Systems Engineering Group, University of York.
Professor	Thomas	McInish	University of Memphis
Professor	Antonio	Mele	University of Lugano and Swiss Finance Institute
Professor	Albert	Menkveld	VU University of Amsterdam
Dr	Yuval	Millo	The London School of Economics
Professor	Gautam	Mitra	Brunel University; CEO OptiRisk Systems
Dr	Richard	Olsen	Chairman and Chief Executive Officer, Olsen Ltd
Dr	Juan Pablo	Pardo-Guerra	The London School of Economics
Professor	Andreas	Park	University of Toronto
Professor	David	Parkes	Harvard University
Professor	Christine	Parlour	University of California, Berkeley
Professor	Richard	Payne	Cass Business School, City University, London
	Anders	Pelli	
Dr	Steve	Phelps	University of Essex
Dr	Alex	Preda	University of Edinburgh

Professor	Tom	Rodden	University of Nottingham
Professor	Bernd	Rosenow	University of Leipzig
Professor	Leonard	Rosenthal	Bentley University
Professor	Gideon	Saar	Cornell University
Professor	Patrik	Sandås	University of Virginia
Professor	Alexander	Schied	University of Mannheim
Dr	Kevin	Sheppard	University of Oxford
Professor	Hyun Song	Shin	Princeton University
Professor	Spyros	Skouras	Athens University
Professor	Didier	Sornette	Chair of Entrepreneurial Risks, ETH Zurich
Professor	Chester	Spatt	Carnegie Mellon University
Dr	Sasha	Stoikov	Head of Research, Cornell Financial Engineering Manhattan
Professor	Martyn	Thomas CBE	Martyn Thomas Associates Ltd
Professor	Ian	Tonks	University of Bath
Professor	Edward	Tsang	University of Essex
Dr	Dimitrios	Tsomocos	University of Oxford
	Nick	Vanston	Experian Decision Analytics
Professor	Daniel	Weaver	Rutgers, Ccpd/Technical Training
Professor	Ingrid	Werner	Ohio State University
Dr	Alexander	Wissner-Gross	Institute Fellow at the Harvard University Institute for Applied Computational Science
	Julie	Wu	
Professor	Mao	Ye	University of Illinois
	Carla	Ysusi	Financial Services Authority
	Ilknur	Zer	London School of Economics
	Feng	Zhan	
Professor	Frank	Zhang	Yale University
	Guannan	Zhao	University of Miami

AUTHORS AND CONTRIBUTORS TO THE EVIDENCE BASE – IMPACT ASSESSMENTS

Professor	Lucy	Ackert	Kennesaw State University
Professor	Amber	Anand	Syracuse University
Professor	Marco	Avellaneda	New York University
Professor	Jean-Philippe	Bouchaud	Chairman of Capital Fund Management; Professor, Ecole Polytechnique
Dr	Paul	Brewer	Economic and Financial Technology Consulting
Professor	Charles	Cao	Pennsylvania State University
Professor	Tarun	Chordia	Emory University
Professor	Kee	Chung	Department Chair, Finance and Managerial Economics, University of Buffalo
Professor	Hans	Degryse	Tilburg University
Professor	Lawrence	Glosten	Columbia University
	Anton	Golub	Manchester Business School
Professor	Roger	Huang	University of Notre Dame
Professor	Sebastian	Jaimungal	University of Toronto; President and CEO, Tyrico Financial Services Inc.
Dr	Elvis	Jarnecic	University of Sydney
Professor	Kenneth	Kavajecz	University of Wisconsin
Professor	John	Keane	University of Manchester
Professor	Yong	Kim	University of Cincinnati
Professor	Paul	Kofman	University of Melbourne
Professor	Bruce	Lehmann	University of California, San Diego
Dr	Wai-Man	Liu	Australian National University
Dr	Marco	Lutat	University of Frankfurt
Professor	Rosario	Mantegna	University of Palermo

	Patrice	Muller	Partner, London Economics Ltd
Professor	Gulnur	Muradoglu	City University, London
Professor	Marco	Pagano	University of Naples Federico II
Professor	Marios	Panayides	University of Pittsburgh
Professor	Ser-Huang	Poon	The University of Manchester
Professor	Ryan	Riordan	Karlsruhe Institute of Technology
Professor	Ailsa	Roell	Columbia University
Professor	Andriy	Shkilko	Wilfrid Laurier University
Professor	Avanidhar	Subrahmanyam	University of California, Los Angeles
Professor	Erik	Theissen	University of Mannheim
	John	Willman	Editorial Consultant
Professor	Jimmy	Yang	Oregon State University

CO-AUTHORS AND CONTRIBUTORS TO THE EVIDENCE BASE

Professor	Ashok	Banerjee	Financial Research and Trading Lab, IIM Calcutta
Professor	Peter	Bossaerts	California Institute of Technology
Dr	Dan	Brown	Entrepreneur in Residence, University College London
Professor	Jaksa	Cvitanic	California Institute of Technology
	Dan	di Bartolomeo	President and founder, Northfield Information Services
Professor	Clemens	Fuest	Oxford University Centre for Business Taxation
Dr	Stefan	Hunt	Financial Services Authority
Dr	Paul	Irvine	University of Georgia
Professor	Donald	MacKenzie	University of Edinburgh
Professor	Linda	Northrop	Carnegie-Mellon University
Professor	Maureen	O'Hara	Cornell University
Professor	Aaron	Pitluck	Illinois State University
Professor	Charles	Plott	California Institute of Technology
Professor	Andy	Puckett	University of Tennessee
	Charlotte	Szostek	University of Bristol
Professor	Philip	Treleaven	University College London
	John	Vella	Oxford University Centre for Business Taxation
Professor	Kumar	Venkataraman	Southern Methodist University
	Susan	von der Becke	ETH Zurich
Dr	Anne	Wetherilt	Senior Manager, Payments and Infrastructure Division, Bank of England
	Xiang	Yu	Brunel University

AUTHORS AND CONTRIBUTORS TO THE EVIDENCE BASE – ATTENDED LEG MEETINGS

	Martin	Abbott	CEO, London Metals Exchange
	Fod	Barnes	Senior Adviser, Oxera
	David	Bodanis	PA Consulting
	Rodrigo	Buenaventura	Head of Markets Participants Division, ESMA
Dr	Alvaro	Cartea	Oxera
	Edward	Corcoran	Her Majesty's Treasury
Dr	Luis Correia	da Silva	Oxera
Dr	Vince	Darley	Head of Analytics and Optimisation, Ocado
	Laurent	Degabriel	Head of Investment and Reporting Division,ESMA
	Simon	Friend	Financial Services Authority
Dr	Martina	Garcia	Her Majesty's Treasury
	Bob	Giffords	Independent analyst and consultant for European banking
	Alasdair	Haynes	Chi-X Europe Limited
	Nichola	Hunter	Head of EBS product management, ICAP
Dr	Andrei	Kirilenko	Commodity Futures Trading Commission
	Shrey	Kohli	London Stock Exchange
	David	Lawton	Financial Services Authority
Dr	Charles-Albert	Lehalle	Head of Quant Research, Cheuvreux Credit Agricole

	Remco	Lenterman	EPTA
	Roger	Liddell	Member, CFTC Global Markets Advisory Committee
	Graham	Lloyd	PA Consulting
	Ted	Moynihan	Oliver Wyman
	John	Romeo	Oliver Wyman
	Tim	Rowe	Financial Services Authority
	John	Schoen	Global Head of EBS Spot Platform and Market Risk, ICAP
	Tom	Stenhouse	London Stock Exchange
	Kee-Meng	Tan	Knight Equity Markets Europe
	Steven	Travers	Head of Regulatory Law and Strategy, London Stock Exchange
	Samad	Uddin	International Organization of Securities Commissions
	Reinder	Van Dijk	Oxera
	Ken	Yeadon	Managing Partner and Co-Founder, Thematic Capital Partners

AUTHORS AND CONTRIBUTORS TO THE EVIDENCE BASE – WORKSHOPS: SCOPING WORKSHOP

Professor	David	Arrowsmith	Queen Mary College, University of London
	Richard	Balarkas	CEO, Instinet Group Incorporated
Dr	Robert	Barnes	UBS
	Graeme	Brister	Legal First
Dr	Mario	Campolargo	European Commission
Dr	Christopher	Clack	University College London
	Robert	Crane	Goldman Sachs
	Rod	Garratt	University of California, Santa Barbara
Dr	Kerr	Hatrick	Deutsche Bank
	Rebecca	Hodgson	Facilitation Team
Professor	Julian	Hunt FRS	Global Systems Dynamics
	Karl	Massey	Deutsche Bank
Professor	Philip	Molyneux	Bangor University
	Paul	Ormerod	Volterra Consulting
	Bill	Sharpe	Facilitation Team
Dr	Christopher	Sier	Technology Strategy Board, Knowledge Transfer Network
	Michael	Straughan	Financial Services Authority
	Stuart	Theakston	GLC Ltd
	Natan	Tiefenbrun	Turquoise
	Nigel	Walker	Technology Strategy Board, Knowledge Transfer Network

WORKSHOP ON DRIVERS OF CHANGE

Dr	Abhisek	Banerjee	The London School of Economics
Dr	Christian	Brownlees	Universitat Pompeu Fabra
	Svetlana	Bryzgalova	The London School of Economics
	Ziad	Daoud	The London School of Economics
Professor	Frank	De Jong	Tilburg University
	Eric	Hunsader	Founder and CEO, Nanex
	Bonsoo	Koo	The London School of Economics
Professor	Albert	Kyle	University of Maryland
Professor	Otto	Loistl	Vienna University of Economics and Business
	Katrin	Migliorati	University of Bergamo
Dr	Sujin	Park	The London School of Economics
	Philippe	Vandenbroeck	Consultant
	Claire	Ward	Financial Services Authority
	Yang	Yan	The London School of Economics

WORKSHOP ON COMPLEX SYSTEMS

Dr	Alfonso	Dufour	University of Reading
Dr	Graham	Fletcher	Cranfield University
	Susan	Gay	Financial Services Authority
Dr	Enrico	Gerding	University of Southampton
Professor	Iain	Main	University of Edinburgh
Professor	Sheri	Markose	University of Essex
Professor	Robert	May	University of Oxford, Workshop Speaker
	Edgar	Peters	First Quadrant
	Chris	Sims	Ignis Asset Management
Dr	Ryan	Woodard	ETH, Zurich, Workshop Speaker

INDUSTRY WORKSHOP, LONDON

Professor	Neil	Allan	Systemic Consult Ltd
	Richard	Anthony	HSBC
Dr	Valerie	Bannert-Thurner	European Operations, FTEN
	Matthew	Beddall	Winton Capital
	Aileen	Daly	Morgan Stanley Europe
	Adrian	Farnham	Turquoise MTF
Dr	Rami	Habib	Algo Engineering
	Peter	Hafez	RavenPack Spain
	Mark	Harding	Head of EMEA Equity Dealing, Invesco
	Lee	Hodgkinson	SVP, Head of European Relationship Management, NYSE Euronext
	Ashok	Krishnan	Bank of America Merrill Lynch
	Miles	Kumaresan	Algonetix UK
	Mike	Lovatt	Accenture
	Dan	Lyons	RGM Advisors
Dr	Mridul	Mehta	Citadel Execution Services Europe
	Paul	Milsom	Legal & General plc
	Alasdair	Moore	Business Development Director and Founder, Fixnetix
	Andrew	Morgan	Deutsche Bank
	Landis	Olson	Hudson River Trading
	Fred	Pedersen	Business Development Manager, Vincorex
	Mike	Powell	Thomson Reuters
	Peter	Randall	Equiduct Europe
Dr	Riccardo	Rebonato	RBS Global Banking & Markets
	Antoine	Shagoury	London Stock Exchange
	David	Shrimpton	Goldman Sachs
	Murray	Steel	Man AHL
Dr	Alex	Suesserott	RTS Realtime Systems
	David	Surkov	Egham Capital Europe
	Adam	Toms	Co-Global Head of Electronic Sales, Nomura Europe
	Laetitia	Visconti	Barclays Capital

INDUSTRY WORKSHOP, SINGAPORE

	Robert	Bilsland	Tullett Prebon
	Tim	Cartledge	Barclays Capital
	Tze Ching	Chang	Capital Markets Department
	Emma	Chen	Prime Minister's Office
	Eng Yeow	Cheu	A*STAR
	Sutat	Chew	Singapore Exchange
	Wei Min	Chin	Accenture
	Bill	Chow	Hong Kong Exchanges and Clearing Limited
	Michael	Coleman	Aisling Analytics

	Petar	Dimitrov	RSR Capital
	Jerome	Guerville	BNPP
	Masayuki	Kato	Tokyo Stock Exchange
	Eugene	Khaimson	Deutsche Bank
	Abhishek	Mohla	UBS
	Jonathan	Ng	Prime Minister's Office
	Richard	Norris	Nomura
	Ivan	Nosha	NOVATTE
	Julian	Ragless	Hong Kong Exchanges and Clearing Limited
	Sunarto	Rahardja	Chi-East
Dr	Sintiani Dewi	Teddy	A*STAR
	Peter	Tierney	NYSE Euronext
	Yeong	Wai Cheong	Algorithms Trading and Research
	Rebecca	Weinrauch	GETCO

INDUSTRY WORKSHOP, NEW YORK

	Paul	Adcock	NYSE Euronext
	Irene	Aldridge	Able Alpha Trading
	Lloyd	Altmann	Accenture
	Scott	Atwell	American Century Investments
	Joe	Bracco	BATS Trading
	Thomas	Chippas	Barclays Capital
	Sanjoy	Choudhury	Nomura
Dr	Ian	Domowitz	Investment Technology Group
	Joseph	Gawronski	Rosenblatt Securities
	Bryan	Harkins	Direct Edge
	Peter	Lankford	Securities Technology Analysis Center
	James	Leman	Westwater Corp.
	Deborah	Mittelman	UBS
	Craig	Mohan	CME Group Inc.
	Chris	Morton	Thomson Reuters
	Adam	Nunes	Hudson River Trading
	Florencia	Panizza	MarketAxess
	John	Prewett	Citigroup
	Adam	Sussman	TABB Group
	Greg	Tusar	Goldman Sachs

WORKSHOP OF CHIEF ECONOMISTS

Dr	Angus	Armstrong	Director of Macroeconomic Research, National Institute of Economic and Social Research
Dr	Riccardo	Barbieri	Chief European economist, Mizuho International
	Malcolm	Barr	Economist, JPMorgan Chase
	Nick	Bate	Economist, Bank of America Merrill Lynch
	Roger	Bootle	Economist, Capital Economics
	George	Buckley	Chief UK Economist, Deutsche Bank
	Willem	Buiter	Chief Economist, Citigroup
	Robert	Carnell	Chief International Economist, ING
Dr	Jonathan	Cave	Senior Research Fellow, RAND (Europe)
	James	Derbyshire	Senior Analyst, RAND (Europe)
	Peter	Dixon	Global Equities Economist, Commerzbank AG
	Patrick	Foley	Chief Economist, Lloyds Banking Group
	Robert	Gardner	Chief Economist, Nationwide Building Society
	Jonathan	Gershlick	Department for Business, Innovation and Skills
Dr	Larry	Hatheway	Chief Economist, UBS
Dr	Andrew	Hilton OBE	Director, Centre for Study of Financial Innovation
Dr	Haizhou	Huang	Chief Strategist, China International Capital Corporation

Dr	Brinda	Jagirdar	General Manager and Head of Economic Research, State Bank of India
Dr	Stephen	King	Group Chief Economist, HSBC
Dr	Eduardo	Loyo	Chief Economist, BTG Pactual
Dr	Gerard	Lyons	Chief Economist, Standard Chartered
Dr	Andrew	McLaughlin	Chief Economist, The Royal Bank of Scotland Group
Professor	David	Miles	Monetary Policy Committee, Bank of England
	Barry	Naisbitt	Chief Economist, Santander UK
Dr	Jim	O'Neill	Chief Economist, Goldman Sachs
	Mario	Pisani	Her Majesty's Treasury
Professor	Vicky	Pryce	Senior Managing Director, FTI Consulting
Dr	Richard	Roberts	SME Economist, Barclays
	Des	Supple	Head of Fixed Income Research, Nomura International
	David	Tinsley	Chief UK Economist, BNP Paribas
	Yan	Wang	Bank of China
Dr	Martin	Weale	Monetary Policy Committee, Bank of England
Dr	Ohid	Yaqub	RAND Europe

ADDITIONAL AUTHORS AND CONTRIBUTORS TO THE EVIDENCE BASE – WORKING PAPER (not mentioned elsewhere)

Professor	Michael	Kearns	University of Pennsylvania
Professor	Fritz	Henglein	University of Copenhagen

ADDITIONAL CONTRIBUTORS – GOVERNMENT OFFICE FOR SCIENCE

	Michael	Reilly	Foresight Research and Knowledge Manager
	Alun	Rhydderch	Foresight International Futures Manager
	Karen	Wood	Foresight Communications Manager
	Martin	Glasspool	Foresight Central Team
Dr	Jane	Jackson	Foresight Central Team

The Government Office for Science would also like to thank the 104 anonymous interviewees who participated in a survey of end users undertaken by Oliver Wyman[3]; the anonymous computer-based traders who presented their views during interviews conducted by Beunza, Millo and Pardo-Guerra[4]; and the academics and industry practitioners who contributed to studies by London Economics[5], and Oxera[6].

3 SR1 (Annex D refers).

4 IN1 (Annex D refers).

5 EIA22 (Annex D refers).

6 EIA21 (Annex D refers).

Annex B References

Abad, D. and Pascual, R. (2007) Switching to a temporary call auction in times of high uncertainty. Available: http://www.cnmv.es/DocPortal/Publicaciones/MONOGRAFIAS/Mon2007_19en.pdf Accessed: 22 August 2012.

Abolafia, M. (2001) Making markets: opportunism and restraint on Wall Street. *Cambridge: Harvard University Press.*

Admati, A.R, DeMarzo P.M., Hellwig, M.F. and Pfleiderer, P. (2010) Fallacies, irrelevant facts, and myths in the discussion of capital regulation: why bank equity is not expensive. Available: https://gsbapps.stanford.edu/researchpapers/library/RP2065R1&86.pdf Accessed: 07 September 2012.

Admati, A.R, DeMarzo P. M., Hellwig, M.F. and Pfleiderer, P. (2012) Debt overhang and capital regulation. Available: http://www.gsb.stanford.edu/news/packages/PDF/AdmatiDebt032612.pdf Accessed: 07 September 2012.

AFM (Netherlands Authority for Financial Markets), (2010) High frequency trading: the application of advanced trading technology in the European marketplace. Available: http://www.afm.nl/~/media/files/rapport/2010/hft-report-engels.ashx Accessed: 22 August 2012.

Amihud, Y., Mendelsohn, H. and Lauterbach, B. (1997) Market microstructures and securities values: evidence from the Tel Aviv Stock Exchange. *Journal of Financial Economics,* **45**(3): 365–390.

Anand, A., McCormick, T. and Serban, L. (2011) Does the make-take structure dominate the traditional structure? Evidence from the options markets (27 June 2011) *Social Science Research Network.* Available: http://ssrn.com/abstract=2074318 Accessed: 22 August 2012.

Angel, J. (1994) Tick size, share prices, and stock splits. Working Paper, Georgetown University. *Journal of Financial Abstracts*, **1**(30). Available: http://bubl.ac.uk/archive/journals/jfa/v01n3094.htm *Social Science Research Network.* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=5486 Accessed: 16 August 2012.

Angel, J., Harris, L. and Spatt, C. (2010) Equity trading in the 21st century. *Social Science Research Network.* Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1584026 http://dx.doi.org/10.2139/ssrn.1584026 http://www.sec.gov/comments/s7-02-10/s70210-54.pdf http://1.usa.gov/LZJf10) Accessed: 16 August 2012.

Antony, J., Bijlsma, M., Elbourne, A., Lever, M. and Zwart, G. (2012) Financial transaction tax: review and assessment. *CPB Discussion Paper 202*, *CPB Netherlands Bureau for Economic Policy Analysis.* Available: http://www.cpb.nl/sites/default/files/publicaties/download/discussion-paper-202-financial-transaction-tax-review-and-assessment.pdf Accessed: 10 September 2012.

Arnuk, S. and Saluzzi, J. (2012) Broken Markets: How High Frequency Trading and Predatory Practices on Wall Street are Destroying Investor Confidence and Your Portfolio. *Financial Times Press.*

BATS Trading Limited. (2009) Pan European Tick Size Pilot, An analysis of results. Available: http://www.batstrading.co.uk/resources/participant_resources/BATSEuro_Tick_Size_Paper.pdf Accessed: 22 August 2012.

Berkman, H., Boyle, G. and Frino, A. (2011) Maker-taker exchange fees and market liquidity: evidence from a natural experiment. *Queenstown, New Zealand: 2011 Financial Management Association International (FMA) Asian conference*, (6–8 April 2011). Available: http://www.fma.org/NewZealand/Papers/GlennBoylePaper.pdf Accessed: 17 September 2012.

Biais, B., Foucault, T. and Moinas, S. (2011) Equilibrium high frequency trading. Available:
http://www.wsuc3m.com/biais%20foucault%20moinas%2014-sept-2011.pdf
https://studies2.hec.fr/jahia/webdav/site/hec/shared/sites/foucault/acces_anonyme/Document/Working%20papers/biais_foucault_moinas_20_oct_2011.pdf
Accessed: 22 August 2012.

Bloomfield, R. and O'Hara, M. (1999) Market transparency: who wins and who loses? *Review of Financial Studies* **12**(1): 5–35 DOI:10.1093/rfs/12.1.5.

Boehmer, E., Fong, K. and Wu, J. (2012) International evidence on algorithmic trading: Edhec Working Paper. *Social Science Research Network.* Available:
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2022034
Accessed: 22 August 2012.

Bommel, J. van. and Hoffmann, P. (2011) Transparency and ending times of call auctions: a comparison of Euronext and Xetra. *LSF Research Working Paper Series* No. 11-09.

Boston Consulting Group (2011) US Securities & Exchange Commission, Organizational Study and Reform. *Boston Consulting Group report*, 10 March 2011. Available:
http://www.sec.gov/news/studies/2011/967study.pdf
Accessed: 23 August 2012.

Brogaard, J. (2010) High frequency trading and its impact on market quality. *Social Science Research Network*. Available:
http://ssrn.com/abstract=1641387
http://www.fsa.gov.uk/static/FsaWeb/Shared/Documents/pubs/consumer-research/jonathan-brogaard-hft.pdf
Accessed: 23 August 2012.

Brogaard, J., Hendershott, T. and Riordan R. (2012) High frequency trading and price discovery. *Social Science Research Network.* Available:
http://faculty.haas.berkeley.edu/hender/HFT-PD.pdf
http://ssrn.com/abstract=1928510 or http://dx.doi.org/10.2139/ssrn.1928510
Accessed: 23 August 2012)

Brunnermeier, M. and Pedersen, L. (2009) Market liquidity and funding liquidity. *Review of Financial Studies*, **22**(6): 2201–2238.DOI:10.1093/rfs/hhn098

Buthe, T. and Mattli, W. (2011) The new global rulers, the privatization of regulation in the world economy. *Princeton University Press*, ISBN 978-0-691-14479-5.

Cartellieri, U. and Greenspan, A. (1996) Global risk management. *William Taylor Lecture Series* ISBN: 1-56708-098-7.

Carlsson, H. and Van Damme, E. (1993) Global games and equilibrium selection. *Econometrica*, **61**(5): 989–1018.

Castura, J., Litzenberger, R. Gorelick, R. and Dwivedi, Y. (2010) Market efficiency and microstructure evolution in U.S. equity markets: a high-frequency perspective. *Working paper*, RGM Advisors, LLC. Available:
http://www.rgmadvisors.com/docs/MarketEfficiencyStudyOct2010.pdf
Accessed: 16 August 2012

Cave, T. (2010) Is the party ending for high-frequency traders? *Financial News* 18 October 2010. Available:
http://www.efinancialnews.com/story/2010-10-18/high-frequency-party-over
Accessed: 16 August 2012

Chaboud, A., Chiquoine, B., Hjalmarsson, E. and Vega, C. (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.* http://www.federalreserve.gov/pubs/ifdp/2009/980/ifdp980.pdf
Social Science Research Network. Available:
http://ssrn.com/abstract=1501135
Accessed: 16 August 2012.

Chicago Board Options Exchange Rulebook (CBOE). Available:
http://cchwallstreet.com/CBOE/Rules
Accessed: 16 August 2012.

Cliff, D. (2009) ZIP60: Further explorations in the evolutionary design of trader agents and online auction–market mechanisms. *IEEE Transactions on Evolutionary Computation*, **13**(1): 3–18.

Comerton-Forde, C., Hendershott, T., Jones, C.M., Moulton, P.C. and Seasholes, M.S. (2010) Time variation in liquidity: the role of market maker inventories and revenues. *Journal of Finance*, **65**(1): 295–331.

Crockett, A. (1998) Banking supervision and financial stability. *William Taylor Memorial Lecture.* Available: http://www.bis.org/speeches/sp981022.htm
Accessed: 22 August 2012.

Danielsson, J. and Shin, H.S. (2003) Endogenous risk, modern risk management: a history. London: Risk Books.

Danielsson, J., Shin, H.S., and Zigrand, J.-P. (2004) The impact of risk regulation on price dynamics, *Journal of Banking & Finance,* **28**(5): 1069–1087.

Danielsson, J., Shin, H.S. and Zigrand, J.-P. (2011) Balance sheet capacity and endogenous risk. *Mimeo* (August 2011). Available:
http://www.riskresearch.org/files/JD-HS-JZ-39.pdf http://www.princeton.edu/~hsshin/www/balancesheetcapacity.pdf
Accessed: 22 August 2012.

Das, R., Hanson, J., Kephart, J. and Tesauro G. (2001) Agent-human interactions in the continuous double auction, *Proceedings of IJCAI-01*, (2): 1169–1176. Available:
http://ijcai.org/Past%20Proceedings/IJCAI-2001/PDF/IJCAI-2001-m.pdf/IJCAI-2001-m.pdf
http://www.research.ibm.com/infoecon/paps/AgentHuman.pdf
Accessed: 22 August 2012.

Degryse, H.A., Jong, F.C.J.M. de and Kervel, V.L. van. (2011) The impact of dark and visible fragmentation on market quality. *Social Science Research Network*. Available:
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1816434
Accessed: 10 September 2012.

Duffie, D. (2010) Presidential address: Asset price dynamics with slow-moving capital. *Journal of Finance*, **65**(4): 1237–1267. DOI: 10.3905/jot.2011.6.2.008.

Easley, D., Lopez de Prado, M. and O'Hara, M. (2011a) The microstructure of the 'Flash Crash': flow toxicity, liquidity crashes and the probability of informed trading. *The Journal of Portfolio Management*, **37**(2): 118–128.

Easley, D., Lopez de Prado, M. and O'Hara, M. (2011b) The exchange of flow toxicity. *Journal of Trading*, **6** (2): 8-13.

Economist (The). (2012) Fine and Punishment. *The Economist* (21 July 2012). Available:
http://www.economist.com/node/21559315
Accessed: 22 August 2012.

Ende, B. and Lutat, M. (2010) Trade-throughs in European cross-traded equities after transaction costs-empirical evidence for the EURO STOXX 50. *2nd International Conference: The Industrial Organisation of Securities Markets: Competition, Liquidity and Network Externalities;* Frankfurt. Available:
https://www.econstor.eu/dspace/bitstream/10419/43270/1/63950793X.pdf
Accessed: 10 September 2012.

European Commission (2004). Directive 2004/39/EC of the European Parliament and of the Council of 21 April 2004 on markets in financial instruments. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32004L0039:EN:HTML
Accessed: 25 September 2012

European Commission (2010). Public Consultation, Review of the Markets in Financial Instruments Directive (MiFID). Available:
http://ec.europa.eu/clima/consultations/0013/consultation_paper_en.pdf
Accessed: 27 September 2012.

European Commission (2011a) Staff Working Paper Impact Assessment (SEC(2011) 1217 final) accompanying the documents: *Proposal for a Regulation of the European Parliament and of the Council on Insider Dealing and Market Manipulation (market abuse)* and *Proposal for a Directive of the European Parliament and of the Council on criminal sanctions for Insider Dealing and Market Manipulation* p 15. Available:
http://ec.europa.eu/internal_market/securities/docs/abuse/SEC_2011_1217_en.pdf
Accessed: 22 August 2012.

European Commission. (2011b) Staff Working Paper, Impact Assessment (SEC(2011) 1226 final) accompanying the documents. *Proposal for a Directive of the European Parliament and of the Council on Markets in Financial Instruments [Recast]* and *Proposal for a Regulation of the European Parliament and of the Council on Markets in Financial Instruments.* Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=SEC:2011:1226:FIN:EN:PDF Accessed: 22 August 2012.

European Commission Taxation and Customs Union. (2012) Taxation of the financial sector. Available: http://ec.europa.eu/taxation_customs/taxation/other_taxes/financial_sector/index_en.htm. Accessed: 10 September 2012.

European Securities and Markets Authority (2011). Consultation Paper (Reference ESMA/2011/224). Guidelines on systems and controls in a highly automated trading environment for trading platforms, investment firms and competent authorities. Available: http://www.esma.europa.eu/system/files/2011-224.pdf Accessed: 1 October 2012.

European Union, The Council of the, Economic and Financial Affairs (ECOFIN). (2012) *Press Release 3178th Council meeting* (22 June 2012). Available: http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ecofin/131141.pdf Accessed: 10 September 2012.

Financial Stability Board. (2012) A Global Legal Entity Identifier for Financial Markets. Available: http://www.financialstabilityboard.rg/publications/r_120608.pdf Accessed: 25 September 2012

Finansinspektionen (Swedish Government). (2012) Investigation into high frequency and algorithmic trading. Available: http://www.fi.se/upload/90_English/20_Publications/10_Reports/2012/htf_eng.pdf Accessed: 22 August 2012.

Fischer, S. (2002) Basel II: Risk management and Implications for banking in emerging market countries. *William Taylor Memorial Lecture.* Available: http://www.iie.com/Fischer/pdf/fischer091902.pdf Accessed: 22 August 2012.

Foucault, T. and Menkveld, A.J. (2008) Competition for order flow and smart order routing systems. *Journal of Finance,* **63**(1): 119–158.

Guardian, The. (2011) Osborne wins wide support in fight against Tobin tax. Available: http://www.guardian.co.uk/business/2011/nov/08/tobin-tax-osborne-eu-clash Accessed: 10 September 2012.

Gennotte, G. and Leland, H. (1990) Market liquidity, hedging, and crashes. *American Economic Review,* **80**(5): 999–1021.

Gomber, P., Arndt, B., Lutat, M. and Uhle, T. (2011) High frequency trading. Technical Report, *Goethe Universität & Deutsche Börse, Social Science Research Network.* Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1858626 Accessed: 22 August 2012.

Gomber, P., Pujol, G. and Wranik, A. (2012) Best execution implementation and broker policies in fragmented European equity markets. *International Review of Business Research Papers,* **8**(2): 144–162.

Griffith-Jones, S. and Persaud, A. Financial Transaction Taxes (2012). Available: http://policydialogue.org/files/publications/FINANCIAL_TRANSACTION_TAXES_Griffith_Jones_and_Persaud_February_2012_REVISED.pdf Accessed: 10 September 2012.

Grossklags, J. and Schmidt, C. (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.

Grossman, S. (1976) On the efficiency of competitive stock markets where trades have diverse information, *Journal of Finance,* **31**(2): 573–585.

Grossman, S. and Stiglitz, J. (1980) On the impossibility of informationally efficient markets. *American Economic Review,* **70**(3): 393–408.

Haldane, A.G. (2009) *Rethinking the financial network.* Available:
http://www.bankofengland.co.uk/publications/speeches/2009/speech386.pdf
Accessed: 22 August 2012.

Haldane, A.G. (2011) *The race to zero.* Available:
http://www.bankofengland.co.uk/publications/speeches/2011/speech509.pdf
Accessed: 22 August 2012.

Haldane, A.G. and May, R.M. (2011) Systemic risk in banking ecosystems. *Nature,* **469**(7330): 351–355.

Halpern, J. and Moses, Y. (1990) Knowledge and common knowledge in a distributed environment. *Communications of the ACM*, **37**(3): 549–587.

Hanke, M., Huber, J., Kirchler, M., and Sutter, M. (2010) The economic consequences of a Tobin Tax – An experimental analysis. *Journal of Economic Behaviour and Organisation*, **74** (1): 58-71.

Harris, L.E. (1997) Order exposure and parasitic traders. Available:
http://www-rct.usc.edu/~/harris/ACROBAT/Exposure.pdf
Accessed: 25 September 2012.

Hasbrouck, J. (2010) Trading costs: a first look. Available:
http://pages.stern.nyu.edu/~jhasbrou/Teaching/2010%20Winter%20Markets/PDFHandouts/Trading%20costs.pdf
Accessed: 15 August 2012.

Hasbrouck, J. and Saar, G. (2011) Low-latency trading, Working paper. *Cornell University, Social Science Research Network.* Available:
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1695460
Accessed: 15 August 2012.

Hemmelgarn, T. and Nicodeme, G. (2010) The 2008 Financial Crisis and Taxation Policy. *EU Commission Taxation Papers, Working Paper No. 20, 2010.* Available:
http://ec.europa.eu/taxation_customs/resources/documents/taxation/gen_info/economic_analysis/tax_papers/taxation_paper_20_en.pdf
Accessed: 10 September 2012.

Hendershott, T., Jones, C. and Menkveld, A. (2011) Does algorithmic trading improve liquidity. *Journal of Finance*, **66**(1): 1–33.

Hendershott, T. and Menkveld, A. (2011) Price pressures. *Social Science Research Network.* http://ssrn.com/abstract=1411943
Accessed: 15 August 2012.

Hendershott, T. and Riordan, R. (2011) Does algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis (Forthcoming article)*. Available:
http://faculty.haas.berkeley.edu/hender/ATMonitor.pdf
Accessed: 10 September 2012.

Hirshleifer, J. (1971) The private and social value of information and the reward to incentive activity. *American Economic Review*, **61**, 561–574.

Holley, E. (2012) More research needed to prevent HFT abuse – Nomura. *The Trade.* Available:
http://www.thetradenews.com/newsarticle.aspx?id=8645&terms=nomura
Accessed: 18 September 2012.

House of Lords European Union Committee. (2012a) 29th Report of Session 2010–12 Towards a Financial Transaction Tax? Available:
http://www.publications.parliament.uk/pa/ld201012/ldselect/ldeucom/287/287.pdf
Accessed: 10 September 2012.

House of Lords European Union Committee. (2012b) 2nd Report of Session 2012-13 MiFID II: Getting it Right for the City and EU Financial Services Industry. Available:
http://www.publications.parliament.uk/pa/ld201213/ldselect/ldeucom/28/28.pdf
Accessed: 25 September 2012.

Hunsader, E. (2010) Analysis of the 'Flash Crash' Date of Event: 20100506, Complete Text. *Nanex Corp.* Available:
http://www.nanex.net/20100506/FlashCrashAnalysis_CompleteText.html.
Accessed: 18 September 2012.

International Monetary Fund (IMF) (2010) A fair and substantial contribution by the financial sector *Final Report for the G-20* (June 2010). Available:
http://www.imf.org/external/np/seminars/eng/2010/paris/pdf/090110.pdf
Accessed: 10 September 2012.

International Organization of Securities Commission (2011) FR09/11, Regulatory Issues Raised by the Impact of Technological Changes on Market Integrity and Efficiency. *Report of the Technical Committee.* Available:
http://www.iosco.org/library/pubdocs/pdf/IOSCOPD361.pdf
Accessed: 18 September 2012.

ITG (Investment Technology Group). (2011) Trading Patterns, Liquidity, and the Citigroup Split. Available:
http://www.itg.com/news_events/papers/CitiSplit2.pdf.
Accessed: 18 September 2012.

Jones, C.M. (2002) A century of stock market liquidity and trading costs. Working paper, *Social Science Research Network*. Available:
http://ssrn.com/abstract=313681 http://dx.doi.org/10.2139/ssrn.313681
Accessed: 18 September 2012.

Jopson, B. (2012) Amazon 'robo-pricing' sparks fears. *Financial Times* (8 July 2012). Available:
http://www.ft.com/cms/s/0/26c5bb7a-c12f-11e1-8179-00144feabdc0.html#axzz24IAnRMrV.
Accessed: 18 September 2012.

Joshua, C. and Stafford, E. (2007) Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics*, **86** (2): 479–512.

Jovanovic, B. and Menkveld, A. (2011) Middlemen in limit-order markets. *Social Science Research Network.* Available:
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1624329.
Accessed: 18 September 2012.

Kaminska, I. (2009) HFT in Europe. *Financial Times Alphaville Blog.* Available:
http://ftalphaville.ft.com/blog/2009/07/24/63651/high-frequency-trading-in-europe/.
Accessed: 18 September 2012.

Kaminska, I. (2011) Algo trading and the Nymex. *Financial Times Alphaville Blog.* Available:
http://ftalphaville.ft.com/blog/2011/03/04/505021/algo-trading-and-the-nymex/.
Accessed: 18 September 2012.

Kay, J. (2012) Interim review of UK equity markets. Available:
http://www.bis.gov.uk/assets/biscore/business-law/docs/k/12-631-kay-review-of-equity-markets-interim-report.pdf.
Accessed: 18 September 2012.

Kay, J. (2012) The Kay Review of UK Equity Markets and Long-Term Decision Making. *Final Report* (July 2012). Available:
http://www.bis.gov.uk/assets/biscore/business-law/docs/k/12-917-kay-review-of-equity-markets-final-report.pdf.
Accessed: 17 September 2012.

Kearns, M., Kulesza, A. and Nevmyvaka, Y. (2010) Empirical limitations of high frequency trading profitability. *The Journal of Trading,* **5**(4): 50–62. DOI: 10.3905/jot.2010.5.4.050. Available:
http://www.cis.upenn.edu/~mkearns/papers/hft_arxiv.pdf.
Accessed: 18 September 2012.

Khandani, A. and Lo, A. (2007) What happened to the quants in August 2007? *Journal of Investment Management*, **5**(4): 29–78.

Khandani, A. and Lo, A. (2011) What happened to the quants in August 2007? Evidence from factors and transactions data. *Journal of Financial Markets,* **14**(1): 1–46.

Kim, K.A. and Park, J. (2010) Why do price limits exist in stock markets? A manipulation-based explanation. *European Financial Management*, **16**(2): 296–318. Available:
http://dx.doi.org/10.1111/j.1468-036X.2008.00456.x.
Accessed: 18 September 2012.

Kirilenko, A., Kyle, A., Samadi, M. and Tuzun, T. (2011) The Flash Crash: the impact of high frequency trading on an electronic market. *Social Science Research Network.* Available:
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1686004.
Accessed: 18 September 2012.

Kissell, R. and Lie, H. (2011) U.S. exchange auction trends: recent opening and closing auction behavior, and the implications on order management strategies. *Journal of Trading,* **6**(1): 10–30 DOI: 10.3905/jot.2011.6.1.010

Krugman, P. (2009) Rewarding bad actors. *New York Times* (2 August 2009). Available: http://www.nytimes.com/2009/08/03/opinion/03krugman.html?_r=1. Accessed: 18 September 2012.

Kyle, A. P. (1985) Continuous auctions and insider trading. *Econometrica*, **53**: 1315–1335.

Larrymore, N.L. and Murphy, A.J. (2009) Internalization and market quality: an empirical investigation. *The Journal of Financial Research*, **32**(3): 337–363. DOI: 10.1111/j.1475-6803.2009.01253.x.

Lee, C., Mucklow, B. and Ready, M. (1993) Spreads, depths, and the impact of earnings information: an intraday analysis. *Review of Financial Studies*, **6**(2): 345–374. DOI:10.1093/rfs/6.2.345.

Leinweber, D.J. (2009) Nerds on Wall Street: Math, Machines and Wired Markets. *John Wiley Publishers.*

London Stock Exchange Group plc (2011) Response to ESMA Consultation: Guidelines on systems and controls in a highly automated trading environment for trading platforms, investment firms and competent authorities ESMA/2011/224, http://www.londonstockexchange.com/about-the-exchange/regulatory/lsegresponsetoesmaconsultationonsystemsandcontrols.pdf Accessed: 10 September 2012.

Lutat, M. (2010) The Effect of Maker-Taker Pricing on Market Liquidity in Electronic Trading Systems – Empirical Evidence from European Equity Trading. *Social Science Research Network.* Available: http://ssrn.com/abstract=1752843. Accessed: 18 September 2012.

Madhavan, A. (2011) Exchange-Traded Funds, Market Structure and the Flash Crash. Working Paper. *BlackRock, Inc. Social Science Research Network.* Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1932925. Accessed: 18 September 2012.

Malinova, K. and Park, A. (2011) Subsidizing Liquidity: The Impact of Make/Take Fees on Market Quality. *Social Science Research Network.* Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1787110. Accessed: 10 September 2012.

Matheson, T. (2011) Taxing Financial Transactions: Issues and Evidence. *IMF WP/11/54* (IMF, March 2011). Available: http://www.imf.org/external/pubs/ft/wp/2011/wp1154.pdf. Accessed: 10 September 2012.

Mehta, N. High Frequency Firms Tripled Trades Amid Rout, Wedbush Says. *Bloomberg* (12 August 2011). Available: http://www.bloomberg.com/news/2011-08-11/high-frequency-firms-tripled-trading-as-s-p-500-plunged-13-wedbush-says.html. Accessed: 18 September 2012.

Menkveld, A. J. (2012). High Frequency Trading and the New-Market Makers, Working Paper. *Social Science Research Network.* Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1722924. Accessed: 18 September 2012.

Milgrom, P. and Stokey, N. (1982) Information, trade and common knowledge. *Journal of Economic Theory,* **26**(1):17–27.

Mill, John Stuart (1846) A system of logic: ratiocinative and inductive; being a connected view of the principles of evidence and the methods of scientific investigation. *Harper & Brothers*.

Morris, S. and Shin, H. (2002) Global games: theory and applications. In Dewatripont, L. and Turnovsky, S. (eds). *Advances in Economics and Econometrics, the Eighth World Congress*. Cambridge University Press.

Nanex. (2012) Facebook IPO How HFT Caused the Opening Delay, and Later Benefited at the Retail Customer's Expense. Available: http://www.nanex.net/aqck/3099.html. Accessed: 18 September 2012.

O'Hara, M. (1995) Market Microstructure Theory. *Blackwell Publishers*.

O'Hara, M. and Ye, M. (2011) Is market fragmentation harming market quality? *Journal of Financial Economics*, **100**(3): 459–474. *Social Science Research Network.* Available: http://ssrn.com/abstract=1356839.
Accessed: 10 September 2012.

Office of Financial Research. (2012) Annual Report 2012. Available: http://www.treasury.gov/initiatives/wsr/ofr/Documents/OFR_Annual_Report_071912_Final.pdf.
Accessed: 17 September 2012.

Oxera. (2011) What would be the economic impact of the proposed financial transaction tax on the EU? (December 2011). Available: http://www.oxera.com/Publications/Reports/2012/What-would-be-the-economic-impact-of-the-proposed-.aspx.
Accessed: 17 September 2012.

Oxera. (2012) What would be the economic impact on the EU of the proposed financial transaction tax? (June 2012). Available: http://marketing.oxera.com/Eventsbr.nsf/0/bf32bb01f3bb843280257a23005a1e50/$FILE/Oxera%20review%20of%20latest%20FTT%20commentary.pdf.
Accessed: 10 September 2012.

Oxford English Dictionary, Available: http://www.oed.com/view/Entry/203876?rskey=MxYIeu&result=1&isAdvanced=false#eid b.
Accessed: 25 September 2012.

Perez, E. (2011) The Speed Traders. *McGraw-Hill Publishers.*

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

Robb, H. W. (1956) Significance of Company and National Standards to Industrial Management. *Harper & Brothers.*

Rothschild, M. and Stiglitz, J.E. (1970) Increasing risk: I. A definition. *Journal of Economic Theory*, **2**(3): 225–243.

Rubinstein, A. (1989) The electronic mail game: Strategic behavior under almost common knowledge. *American Economic Review*, **79**(3): 385–391.

Samuelson, P.A. (1947) Foundations of economic analysis. *Harvard University Press*.

Schulmeister, S., Schratzenstaller M. and Picek, O. A (2008) General Financial Transaction Tax. Motives, Revenues, Feasibility and Effects. *Research Study WIFO*. (March 2008). Available: http://www.wifo.ac.at/wwa/jsp/index.jsp?typeid=8&display_mode=2&fid=23923&id=31819.
Accessed: 10 September 2012.

Semeta, A. (2012) Financial Transactions Tax: The Way Ahead. *Meeting of Members of the Finance and Fiscal Committees in the Danish Parliament, Copenhagen,* (19 March 2012). Available: http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/12/196.
Accessed: 18 September 2012.

Shin, H.S. (1996) Comparing the robustness of trading systems to higher order uncertainty. *Review of Economic Studies,* **63**(1): 39–59.

Shin, H.S. (2010) Risk and liquidity. *2008 Clarendon Lectures in Finance, Oxford University Press.*

Sloan, S. (2012) High-frequency trading has 'disconnected' public, NYSE CEO says. *Bloomberg* (20 June 2012). Available: http://www.bloomberg.com/news/2012-06-20/high-frequency-trading-has-disconnected-public-nyse-ceo-says.html.
Accessed: 18 September 2012.

Smith, V.L. (1982) Microeconomic systems as an experimental science. *American Economic Review,* **72**(5): 923–955.

Storkenmaier, A. and Wagener, M. (2011) Do we need a European 'National Market System'? Competition, arbitrage, and suboptimal executions. *Social Science Research Network*. Available: http://www.efmaefm.org/0EFMAMEETINGS/EFMA%20ANNUAL%20MEETINGS/2011-Braga/papers/0069_update.pdf.
Accessed: 10 September 2012.

Subrahmanyam, A. (1998) Transaction Taxes and Financial Market Equilibrium. *Journal of Business*, **71**(1): 81–118. Available:
http://www.jstor.org/stable/10.1086/209737.
Accessed: 11 September 2012.

Surowiecki, J. (2011) Turn of the century (April 2011). Available:
http://www.wired.com/wired/archive/10.01/standards_pr.html.
Accessed: 18 September 2012.

TABB Group. (2012) US equities market 2012: mid-year review. Available:
http://www.tabbgroup.com/PublicationDetail.aspx?PublicationID=1129&MenuID=44&ParentMenuID=2&PageID=43.
Accessed: 18 September 2012.

Thurner, S., Doyne Farmer, J. and Geanakoplos, J. (2012) Leverage causes fat tails and clustered volatility. *Quantitative Finance*, **12**(5): 695–707 DOI:10.1080/14697688.2012.674301.

Troianoivski, A. (2012) Networks built on milliseconds. *The Wall Street Journal* (30 May 2012). Available:
http://on.wsj.com/JUMhiH.
Accessed: 16 August 2012.

US Commodity Futures Trading Commission and the US Securities & Exchange Commission. (2010) Findings regarding the market events of May 6, 2010. Available:
http://www.sec.gov/news/studies/2010/marketevents-report.pdf
Accessed: 24 September 2012.

US Commodity Futures Trading Commission. (2010) Glossary. Available:
http://www.cftc.gov/ucm/groups/public/@educationcenter/documents/file/cftcglossary.pdf
Accessed: 1 October 2012.

US Securities & Exchange Commission. (2010) Concept release on equity market structure. *Release No. 34-61458; File No. S7-02-10, page 45.*

US Securities & Exchange Commission. (2012a) SEC approved the NYSE proposed pilot on July 3, 2012. *Release No. 34-67347; File Nos. SR-NYSE-2011-55: SR-NYSE Amex-2011-84*

US Securities & Exchange Commission. (2012b) Self-Regulatory Organizations; The NASDAQ Stock Market LLC; NYSE Arca, Inc. *Order Instituting Proceedings to Determine Whether to Approve or Disapprove Proposed Rule Changes Relating to Market Maker Incentive Programs for Certain Exchange-Traded Products Release No. 34-67411; File Nos. SR-NASDAQ-2012-043: SR-NYSE Arca-2012-37).* Available:
http://www.sec.gov/rules/sro/nasdaq/2012/34-67411.pdf.
Accessed: 18 September 2012.

Vaughan, D. (1997) The Challenger launch decision: risky technology, culture and deviance at NASA. *University of Chicago Press.*

Vella, J., Clemens, F. and Schmidt-Eisenlohr. T. (2011) The EU Commission's Proposal for a Financial Transaction Tax. *British Tax Review* **2011** (6): 607-621

Volcker, P.A. (2011) Three years later: unfinished business in financial reform. *William Taylor Memorial Lecture.* Available:
http://graphics8.nytimes.com/packages/pdf/business/23gret.pdf and
http://www.group30.org/images/PDF/ReportPDFs/GRP30_WTML13_Volcker_FNLlo.pdf.
Accessed: 18 September 2012.

Volcker, P.A. (2012) Unfinished business in financial reform. *International Finance*, **15**(1): 125–135. DOI: 10.1111/j.1468-2362.2011.01294.x.

Weaver, D.G. (2011) Internalization and market quality in a fragmented market structure. *Social Science Research Network.* Available:
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1846470.
Accessed: 18 September 2012.

Zhang, F. (2010) High-frequency trading, stock volatility, and price discovery. *Social Science Research Network.* Available:
http://ssrn.com/abstract=1691679.
Accessed: 18 September 2012.

Zigrand, JP (2005) Rational asset pricing implications from realistic trading frictions. *Journal of Business*, **78**(3): 871–892.

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

Annex C Glossary of terms and acronyms

Terms

The following defines the terms used in this Report. A list of acronyms is given at the end.

Algorithmic trading (AT) – different definitions of this term are used in various parts of the world. For example, the following provides those used by the European Commission Directorate General Internal Market and Services (DG MARKT) and the US Commodity Futures Trading Commission (CFTC). The use of AT in this report is broadly consistent with both of these definitions:

> **DG MARKT:** "Automated trading also known as algorithmic trading can be defined as the use of computer programmes to enter trading orders where the computer algorithm decides on aspects of execution of the order such as the timing, quantity and price of the order."[1]

> **CFTC:** "The use of computer programs for entering trading orders with the computer algorithm initiating orders or placing bids and offers."[2]

Bid-ask spread – the difference in price between the highest price that a buyer is willing to pay for an asset and the lowest price for which a seller is willing to sell it. This difference in price is an implicit trading cost to pay for immediate liquidity.

Broker crossing networks – trading platforms in which traders submit buy and sell orders and trades are executed for matching orders at an exogenous price; typically the midpoint of the current bid and offer spread. As there are often more orders on one side than the other, not all orders execute. Trades that do execute, however, do so at a better price than they would have in the market.

Cloud computing – Ultra large scale data centres (vast warehouses full of interconnected computers) which can be 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.

Computer-based trading – this Project has taken a broad interpretation of computer-based trading (CBT). A useful taxonomy of CBT was proposed in DR5, which identifies four 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; third, they may be classified as engaging in either uninformed or informed trading; and, fourth, the trading strategy can be generated by the algorithm itself or, alternatively, the algorithm is used only to optimally implement a decision taken otherwise.

Dark pool – trading systems in which buy and sell orders are submitted anonymously and are not displayed to the public markets until execution. Broker crossing networks defined above are a type of dark pools.

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.

Exchange traded funds (ETF) – 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.

1 See European Commission (2010).

2 See CFTC (2010).

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

High Frequency Trading (HFT) – different definitions of this new term are used in various parts of the world. For example, the following provides those used by European Securities and Markets Authority (ESMA) and the Securities and Exchange Commission (SEC) in the US. The use of HFT in this report is broadly consistent with both of these definitions:

> **ESMA:** "Trading activities that employ sophisticated, algorithmic technologies to interpret signals from the market and, in response, implement trading strategies that generally involve the high frequency generation of orders and a low latency transmission of these orders to the market. Related trading strategies mostly consist of either quasi market making or arbitraging within very short time horizons. They usually involve the execution of trades on own account (rather than for a client) and positions usually being closed out at the end of the day."[3]
>
> **SEC:** "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)."[4]

Informed or value-motivated traders – aim to profit by trading on the basis of information and 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.

Internalisation – the practice whereby some customer trades are executed internally by brokers or intermediaries and so do not reach public markets.

Inventory traders – aim to profit by merely providing liquidity and act as 'market-makers': they hold a sufficiently large 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. Inventory traders can, in principle, operate profitably without recourse to any information external to the market in which their instruments are being traded.

Layering and spoofing – layering refers to entering hidden orders on one side of the book (for example, a sell) and simultaneously submitting visible orders on the other side of the book (buys). The visible buys orders are intended only to encourage others in the market to believe there is strong price pressure on one side, thereby moving prices up. If this occurs, the hidden sell order executes, and the trader then cancels the visible orders. Similarly, spoofing involves using, and immediately cancelling, limit orders in an attempt to lure traders to raise their own limits, again for the purpose of trading at an artificially inflated price.

Liquidity – the ability to buy or sell an asset without greatly affecting its price. The more liquid the market, the smaller the price impact of sales or purchases.

Locked or crossed market – a locked or crossed market is where the bid-ask spread is zero (negative) and is evidence of poor linkages between markets and investors.

Long-only macro trading – long-only macro trading is where a trader maintains a portfolio of holdings that are bought only in the expectation that their market value will increase (that is, the trader is taking long positions, as opposed to short positions where the trader would benefit if the market value decreases); and where the trader's alterations to the portfolio are driven by macroeconomic factors,

3 See ESMA (2011).

4 See SEC (2010).

such as national interest rates, which tend to alter relatively slowly, rather than the second-by-second fluctuations commonly exploited by high frequency traders.

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

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

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

Market abuse[5] – consists of market manipulation and insider dealing, which could arise from distributing false information, or distorting prices and improper use of insider information.

Market abuse may be grouped into the following seven categories:

Insider dealing: when an insider deals, or tries to deal, on the basis of inside information

Improper disclosure: where an insider improperly discloses inside information to another person

Manipulating transactions: trading, or placing orders to trade, that gives a false or misleading impression of the supply of, or demand for, one or more investments, thus raising the price of the investment to an abnormal or artificial level.

Misuse of information: is behaviour based on information that is not generally available but would affect an investor's decision about the terms on which to deal.

Manipulating devices: refers to trading, or placing orders to trade, employing fictitious devices or any other form of deception or contrivance.

Dissemination: refers to giving out information that conveys a false or misleading impression about an investment or the issuer of an investment where the person doing this knows the information to be false or misleading.

Distortion and misleading behaviour: refers to behaviour that gives a false or misleading impression of either the supply of, or demand for, an investment; or behaviour that otherwise distorts the market in an investment.

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

Momentum ignition – entry of orders or a series of orders intended to start or exacerbate a trend, and to encourage other participants to accelerate or extend the trend in order to create an opportunity to unwind/open a position at a favourable price.

Multilateral trading facility (MTF) – According to MiFID Directive, article 4, no. 15) a. MTF is "a multilateral system, operated by an investment firm or a market operator, which brings together multiple third-party buying and selling interests in financial instruments - in the system and in accordance with non-discretionary rules – in a way that results in a contract in accordance with the provisions of Title II."[6]

Order anticipation strategies – a trader looks for the existence of large (for example) buyers, in the objective of buying before these orders, in order to benefit from their impact.

Order preferencing – arrangements whereby orders are sent to a prespecified exchange or broker/dealer in exchange for a per share payment.

5 The source for definitions of market abuse related terms are EU Commission Staff Working Paper, (2011b) AFM, (Netherlands Authority for Financial Markets), (2010).

6 See European Commission (2004).

Passive limit orders – these provide the counterparty for traders wishing to find a buyer or seller in the market.

Ping orders – entering small orders into a market in order to assess the real level of liquidity on that venue (beyond that displayed).

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

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

Primary market – when a company issues equities (shares) to raise capital, that is the primary market in action.

Quote stuffing – entering large numbers of orders and/or cancellations/updates to orders so as to create uncertainty for other participants, slowing down their process and to camouflage the manipulator's own strategy.

Secondary market – when the shares from the primary market are subsequently traded among investors and speculators, that is the secondary market in action.

Statistical arbitrage – also known as '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 and that the relationship will revert to its mean in due course.

Suspicious transaction report (STR) – reports to competent authorities required under Article 6(9) of the Market Abuse Directive where a person professionally arranging transactions reasonably suspects that a transaction might constitute insider dealing or market manipulation.

The 'touch' – According to the Oxford English Dictionary, the touch is "the difference or spread between the highest buying price and the lowest selling price in a commodities or share market."[7]

Trade-through – when a trade is executed at a price on one venue that is inferior to what was available at another venue at the same time.

Transaction costs – the costs traders incur to buy or sell an asset.

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

7 See Oxford English Dictionary.

Acronyms

AES	automated execution system
AT	algorithmic trading
ATS	Alternative Trading System
BATS	Better Alternative Trading System
CBT	computer-based trading
CDO	collateralised debt obligation
CDS	credit default swaps
CFD	contracts for difference
CFTC	Commodity Futures Trading Commission
CLOB	Central Limit Order Book
CME	Chicago Mercantile Exchange
CoCo	conditional convertible debt instrument
CPU	central processing unit
CRA	credit rating agency
DMA	direct market access
EBS	electronic broking services
EDHEC	Ecole des Hautes Etudes Commerciales
EFDC	European Financial Data Centre
EMEA	Europe, the Middle East and Africa
EOFR	European Office of Financial Research
ESMA	European Securities and Markets Authority
ETF	exchange traded fund
ETP	exchange traded product
EU	European Union
FESE	Federation of European Securities Exchanges
FINRA	US Financial Industry Regulatory Authority
FPC	Financial Policy Committee
FPGA	field programmable gate array
FSA	Financial Services Authority
FSOC	Financial Stability Oversight Council
FX	foreign exchange
GDP	gross domestic product
GETCO	Global Electronic Trading Company
HFT	high frequency trading
HPC	high performance computing
HRO	high reliability organisation
IMF	International Monetary Fund
IOSCO	International Organization of Securities Commissions
IPO	initial public offering
IPR	intellectual property rights
ISE	Istanbul Stock Exchange
ISO	International Organisation for Standards
LEI	legal entity identifier
LIBOR	London Interbank Offered Rate
LSE	London Stock Exchange
LTI	loan to income
LTV	loan to value
MFR	monthly fill ratio
MiFID	Markets in Financial Instruments Directive
MQL	minimum quote lifespan
MTF	multilateral trading facility
NMS	national market system
OBI	order book imbalance
OECD	Organisation for Economic Co-operation and Development
OER	order-to-execution ratio

OFR	Office of Financial Research
OTC	over-the-counter
OTR	order-to-trade ratio
PC	personal computer
PTP	price/time priority
RLP	retail liquidity provider
RoE	return on equity
SAM	shared appreciation mortgages
SEAQ	Stock Exchange Automated Quotation
SEC	Securities and Exchange Commission
SI	systematic internaliser
SMS	standard market size
SOR	smart order routing/router
STP	straight-through processing
STR	suspicious transaction report
TAIFEX	Taiwan Futures Exchange
TRF	trade reporting facility
TWAP	time-weighted average price
UNSCC	United Nations Standards Coordination Committee
VWAP	volume-weighted average price

Annex D Project reports and papers

Executive Summary

Driver Reviews

DR1: What has happened to UK equity market quality in the last decade?

DR2: Feedback effects and changes in the diversity of trading strategies

DR3: Technology trends in the financial markets

DR4: The global financial markets – an ultra-large-scale systems perspective

DR5: Computer-based trading, liquidity and trading costs

DR6: An ecological perspective on the future of computer trading

DR7: Crashes and high frequency trading

DR8: Automated analysis of news to compute market sentiment

DR9: Leverage, forced asset sales and market stability

DR10: High frequency trading information and profits

DR11: Impersonal efficiency and the dangers of a fully automated securities exchange

DR12: High frequency trading and price efficiency

DR13: Studies of interactions between human traders and algorithmic trading systems

DR14: Prospects for large scale financial systems simulation

DR15: Impact of special relativity on securities regulation

DR16: Electronic trading and market structure

DR17: The evolution of algorithmic classes

DR18: Pricing liquidity in electronic markets

DR19: Market fragmentation in Europe: assessment and prospects for market quality

DR20: Computer-based trading and market abuse

DR21: High frequency trading and the execution cost of institutional investors

DR22: High frequency trading and end of day price manipulation

DR23: Algorithmic trading and changes in firms' equity capital

DR24: The impact of high frequency trading on market integration – an empirical examination

DR25: Exploring the 'robot phase transition' in experimental human-algorithmic market

DR26: Computer trading and systemic risk – a nuclear perspective

DR27: Brave new world: quantifying the new instabilities and risks arising in sub second algorithmic trading

DR28: High frequency trading – assessing the impact on market efficiency and integrity

DR29: Systemic risk arising from computer-based trading and connections to the empirical literature on systemic risk

DR30: Trust and reputation in financial services

DR31: Standards in financial services

Final Report

Economic Impact Assessments

- **EIA1:** Minimum resting times and order message-to-trade ratio
- **EIA2:** Minimum resting times and transaction-to-order ratios
- **EIA3:** Minimum resting times, call markets and circuit breakers
- **EIA4:** Circuit breakers
- **EIA6:** Tick sizes
- **EIA7:** Tick sizes
- **EIA8:** Minimum obligations for market makers
- **EIA9:** Circuit breakers
- **EIA10:** Internalisation
- **EIA11:** Continuous market vs. randomised stop auctions and alternative priority rules
- **EIA12:** Maker-taker pricing
- **EIA13:** CLOB vs. exchange order books
- **EIA16:** Algorithmic regulation
- **EIA17:** Market abuse and surveillance
- **EIA18:** Ratio of orders to transactions
- **EIA19:** Minimum obligations for market makers
- **EIA20:** Harmonised circuit breakers
- **EIA21:** Economic impact of MiFID regulation
- **EIA22:** Economic impact assessment of proposal for CBT

Surveys and Interviews

- **SR1:** Survey of end-users
- **IN1:** Interviews with computer-based traders

Workshop Reports

- **WR1:** Chief Economists workshop report
- **WR2:** Industry workshop report
- **WR3:** Drivers of change and market structure
- **SC1:** Consolidated scenarios

Project reports and papers are freely available to download at http://www.bis.gov.uk/foresight

Note 1: Some report numbers were initially allocated but were not subsequently used.
Note 2: All driver reviews and economic impact assessments have passed a double blind peer review process.

Annex E Possible future scenarios for computer-based trading in financial markets

In the future, the development and uptake of computer-based trading (CBT) will be substantially influenced by a range of important drivers of change, acting alone and in concert. Regulation will be particularly important[1], as will technology development and roll-out[2], but other key drivers identified in this Project include[3]: demographic shifts, global economic cycles, geopolitics, changes in riskless assets, changing asset classes, competition and changes in (dis)intermediation. Conversely, CBT will itself have a profound effect on market operation, market structure and, in turn, feed back to influence some of those key drivers of change.

While the time horizon of this Project is just ten years or so, there are considerable uncertainties inherent in many of these drivers of change. For this reason, it was impossible to determine a single 'most probable' future scenario. Therefore, this Project has developed four plausible future scenarios – in essence these sample the future 'possibility space'[4].

While the development of these scenarios has not substantially affected the detailed assessments of individual regulatory measures in Chapter 6, the scenarios generated were considered useful in stimulating thinking about the wider context in which CBT operates. For this reason, the following provides a brief overview of the four scenarios considered.

Four future scenarios

There are many ways in which scenarios for future financial markets can be cast. However, as computer trading was a central concern, two broad drivers of change were considered particularly important following the workshops mentioned above. One driver is the extent to which future growth is sustained and convergent and the other is the extent to which markets are open and reliant on decentralised technologies. These are therefore represented as the horizontal and vertical axes respectively in Figure E.1. The quadrants of this figure therefore define four contrasting scenarios, described briefly below.

1 See Chapter 6.

2 See Chapter 2.

3 These drivers of change were identified in workshops in London, Singapore and New York involving leading industry practitioners, and a workshop of chief economists in London. See Chapter 1 for a brief description of these drivers and the various workshop reports. See WR1, WR2, WR3, SC1 (Annex D refers).

4 Each of the workshops mentioned emphasised different drivers of change as particularly important (see Figure 1.1 in Chapter 1) although regulation and demographic change emerged as consistently important. The scenarios presented here are not therefore representative of any one workshop, but rather an amalgamation of views expressed.

Figure E.1: The four possible future scenarios, illustrated as quadrants spanning a two-dimensional space where one dimension is the extent which future growth is sustained and convergent (horizontal axis) and the other is the extent which markets are open and reliant on decentralized technologies (vertical axis)

Open markets, distributed technologies

Scenario D – Geopolitical and economic competition

- Competing world powers and economic models
- Light touch regulation but low growth
- Investment shifts to emerging markets
- Companies in much of Asia owned by state and financial elite
- In Asia, retail trading of derivatives and other synthetics explodes – copied in the West
- Bank shrinkage and deleveraging
- Real assets preferred to financial
- Shadow banking thrives; households SMEs suffer
- New instruments and online exchanges for company financing

Scenario A – Back to the 1980s

- Globalisation advances
- Global imbalances resolved: US deficit stabilise; China growth stabilises
- Increased financial market activity (finance increases share of GDP)
- Contained EU defaults
- Competition leads to technology innovation, pressure on traditional trading margins
- Tech-literate new trading generation
- New platforms, products
- Interlinked global exchanges, single exchange view
- Competing trading and clearing components marketed to exchanges

Sluggish and fragmented growth

Sustained and convergent growth

Scenario C – Retrenchment and global tensions

- Global lost decade – Economic systems worldwide retrench in the face of severe challenges
- Tensions rise across the globe – war?
- Bank failures, restructurings, nationalisation
- Capital controls, protectionism
- Dollar loses reserve currency status
- Pressure on exchanges, trading firms, leads to consolidation, rebundling, monopolies
- Proliferation of synthetic products with opaque structures, responding to demand for 'safe returns'
- HFT grows. Churning and copycat strategies "are adopted"

Scenario B – Back to the 1950s (post-war boom)

- More regulation, growth remains high
- Emerging economies take the lead
- International institutions play an increasing role in governing economic and political systems
- Rebalancing of capital and trading volumes to emerging markets
- Financial markets reconfigure along regional lines
- Regional exchanges dominate, interconnected in a carefully regulated system
- Less complex assets
- Responding to low beta margins, more macro strategies lead to correlation, lower volumes

Closed markets, centralised technologies

Scenario A – Back to the 1980s

In this scenario, global imbalances begin to be resolved as the savings and investments imbalance between East and West reduces. Globalisation makes steady progress; growth returns to the developed world and picks up again in the emerging economies.

The Japanese economy develops (modest) surpluses, US deficits stabilise, Europe survives a series of 'corrective' defaults without major disruption and China's growth is stabilised by the emergence of healthy levels of domestic demand.

Financial markets account for an increasing proportion of GDP, although equities lose market share to derivatives and other synthetic products. Increasing amounts of investment and trading activity are accounted for by hedge funds, in part to avoid barriers and regulations. New asset classes emerge, such as water futures and country insurance bonds.

Novel platforms and financial services are also developed: for example, software providers sell competing trading and clearing components to exchanges and also offer retail versions of their products, such as algorithmic trading on mobile handsets. Retail investment volumes increase strongly.

The interconnection of global exchanges continues, offering a single view of multiple exchanges and opening the way to a 'financial grid' system. The market for trading intermediaries attracts new entrants; competition leads to technology innovation and pressure on traditional trading margins.

Overall in this scenario liquidity is good, financial stability increases, volatility declines, prices efficiently incorporate all knowledge and risks, transaction costs are low and markets effectively self-regulate against opportunistic or predatory behaviour.

Scenario B – Back to the 1950s (post-war boom)

In this scenario, markets have become less open and finance is increasingly regulated around the world. Capital and trading volumes are 'rebalanced' in favour of emerging markets, which are better able to sustain growth, while Western economies remain encumbered by their debt burdens. International institutions play an increasing role in governing economic and political systems, offering a greater voice to the new economies. The Washington consensus of lightly regulated financial markets and capital flows has passed.

Financial markets are reconfigured along regional lines and investment is biased towards domestic markets. Governments use regulation to capitalise on – and protect – local advantages. Globalisation declines, but does not disappear altogether; trade has reached a plateau; some regional trading blocs (such as the EU) remain effective. Trade deals tend to be bilateral rather than multilateral, and more in the interest of the new economies.

The UK seeks to rediscover the macroeconomic policy stance of the post-war years, when higher levels of financial market regulation coexisted with annual growth rates of 5% or more. Such growth rates prove out of reach in the 2010s however, as debt is paid down.

Protectionist government intervention means financial resources are not allocated efficiently; there is more 'strategic' allocation and less short-term profit maximisation. Price discovery is distorted; risk is not efficiently priced. Savings in the emerging countries are reinvested there; the West has to bid over the odds to attract investment capital. Political support for free trade and open markets weakens in the West, leading to a gradual departure of entrepreneurial talent.

Institutional investment outweighs retail investment; nominal interest rates do not equalise on a global basis. Large regional exchanges dominate; smaller firms and platforms are perceived as risky after well-publicised failures. Stock exchange regulation privileges consumer protection. There are more manufacturing jobs and the financial services industry accounts for a smaller proportion of GDP.

The financial world sees a return to less exotic products and services, epitomised by the return of the building-society model. There is more interest in risk sharing than aggressive (financial) gain seeking. Continued high frequency trading (HFT) use leads to reduced trading margins for institutions, while macro strategies lead to correlation thus reducing activity levels.

As this scenario entails compartmentalisation, asset simplification and financial restrictions, especially of international flows, liquidity and volatility will generally reduce, but financial stability will increase. Markets will not price properly for risk, which will harm price efficiency and discovery and increase the volatility of real returns to specific assets. Unit transaction costs will increase. Market integrity will decline but market insiders and those who circumvent the markets will do well.

Scenario C – Retrenchment and global tensions

In this scenario, the collapse of the Euro and ensuing depression leads to the reintroduction of capital controls in Europe; banks are nationalised amid a resurgence of protectionism. Economic systems worldwide retrench in the face of severe challenges.

The strengthening US dollar loses its reserve currency role as the USA attempts to improve its current account deficits by devaluation (rather than reform of fiscal policy and savings–investment imbalances), triggering inflation and increasing the perceived riskiness of both US sovereign debt and other dollar-denominated assets. The resulting confusion and the lack of an alternative 'risk-free' reference asset, lead to violent and unpredictable price fluctuations.

Globally, there are persistent increases in deficits and a Japanese-style lost decade lies ahead. The West reacts by blaming others and turning protectionist. There are bank failures, restructurings and nationalisations.

The loss of equilibrating global trades and the resulting restrictions on free movement of assets and entry exacerbate inequality within and among nations. This is extended by the failure of markets properly to price claims in different dates and states.

Populist politicians and parties come to the fore on a wave of discontent. 'Arab spring' turns into 'Arab winter'; developed and emerging nations blame each other for their growing economic troubles; China and India fight for access to the remaining world markets. These tensions are set to escalate dangerously.

Exchanges are hit by falling volumes as people lose confidence in markets and seek liquidity. Several major trading firms fail. This leads to concentration; surviving firms use their dominance to control access to exchanges and 're-bundle' elements of the trading and clearing process.

Responding to demand from pension funds and the withdrawal of support from central banks, liquidity provision is increasingly privatised through 'liquidity transfers', spreading liquidity risk around the banking system.

Structured products are used to offer 'safe' returns to the retail market. 'Alpha' trading strategies – which aim to achieve returns that beat a relevant benchmark – require taking on increasing principal risk. HFT grows, as returns are desperately sought. Churning and copycat strategies are adopted, and multiple feedback loops increase the likelihood of a catastrophic financial event.

As in Scenario B, this scenario involves widespread use of capital controls and a strong 'home bias'. In contrast, however, it has limited or no growth. In consequence, although liquidity is reduced, increases in financial stability cannot be taken for granted, because low growth threatens the existence of a riskless asset. On the other hand, short of catastrophe, low growth depresses volatility and assets prices alike. Transaction costs are up. The pervasiveness of inefficient barriers to economic activities produces corruption and other threats to market integrity.

Scenario D – Geopolitical and economic competition

In this scenario, regulation on the whole remains 'light touch', but growth still falters. As stagnation persists, however, some nations opt for greater regulation and restrict access to their financial markets. World powers have entered a phase of geopolitical and economic competition.

Western economies stagnate; quantitative easing has become less effective. Even China sees GDP growth fall, but the 'Asian model' of a state-dominated economy remains in the ascendant. Companies in much of Asia are owned by the state and financial elite – share ownership does not converge with patterns in the West. The euro, sterling and yen lose value, and the US dollar shows signs of weakness. More trade is carried out in renminbi.

The US deleverages, in part by renegotiating or defaulting on debts. Underemployment has lowered productivity growth, while differing skill levels exacerbate wage inequality. Rising inequality has increased political fragmentation and extremism.

In countries that persist with light-touch regulation and open markets, banks and some other financial institutions have shrunk. This is partly a consequence of competition, but also to do with deleveraging. The flight of capital to emerging – and liberalising – markets prolongs the stagnation of the developed economies.

Investors have become desperate for yield. 'Buy and hold' strategies and pension funds underperform. As a result, capital flows from the USA to China, Latin America and other emerging markets, and from traditional financial assets to 'real' investments; some markets struggle to handle these shifts and episodes of illiquidity occur.

The shadow-banking sector does well, but households and small firms struggle for access to capital. New instruments and online exchanges are created for company financing, as traditional capital-raising falters.

Seeking new sources of revenue in a stagnant trading environment, firms in the West offer trading apps to their customers, having seen them take off among Asian retail investors. Retail trading of derivatives and other synthetics grows strongly, though many see this as a betting culture taking hold in the absence of reliable yields from traditional investments.

Problems considered less urgent, such as the environment, take a back seat to short-term economic policy. Domestic and international political tensions are stoked by migration, continued economic malaise and growing differences in economic performance.

Overall, this scenario embodies openness without strong global growth. This exacerbates unstable global financial flows. Liquidity is unreliable, financial stability is threatened by a series of speculative bubbles and crashes and volatility increases. The unstable flight of capital through a fragmented financial system indicates effective price discovery and low transactions costs. The zero-sum nature of this scenario makes the threats to market integrity even worse than in Scenario C.

Printed in the UK on recycled paper with a minimum HMSO score of 75
First published October 2012.
The Government Office for Science