Perspectives on High-Frequency Trading

Fokion Zervoudakis, Donald Lawrence,
George Gontikas, and Majd Al Merey
Department of Computer Science, University College London, London, UK
{f.zervoudakis,d.lawrence,gontikas,malmerey}@cs.ucl.ac.uk

Abstract

High-frequency trading (HFT) is *algorithmic trading* that utilizes sophisticated low latency technologies to conduct automated market making, spread trading, arbitrage and macrotrading. HFT currently generates approximately 35% and 70% of UK and U.S. equity trades, respectively, and thereby exerts significant influence over the global financial markets. The ubiquitous diffusion of HFT systems proceeds in contrast to the discord and controversy engendered by those systems. Financial regulators, market participants, academics and members of the media express diverging, and sometimes exaggerated, views with respect to HFT benefits and risks. The *May 6, 2010 Flash Crash*, which temporarily erased $1 trillion in market value, highlights the potential for amplification of financial anomalies by HFT systems. We argue for an informed dialogue between stakeholders to realize a common domain of discourse, and an appropriate and resilient regulatory framework. To this end, we have conducted a survey that collates the opinions of academics, market participants and regulators, with respect to HFT benefits and risks. Our findings are presented in this report.

1 Introduction

The term *algorithmic trading* describes financial market transactions that are based on instructions generated by quantitative algorithms [1]. These algorithms can be encoded in a programming language, and executed automatically by computers. *High-frequency trading* (HFT) is algorithmic trading that utilizes sophisticated low latency technologies to conduct automated market making, spread trading, arbitrage and macrotrading [2, 3]. We consider low latency environments to be those environments that can support

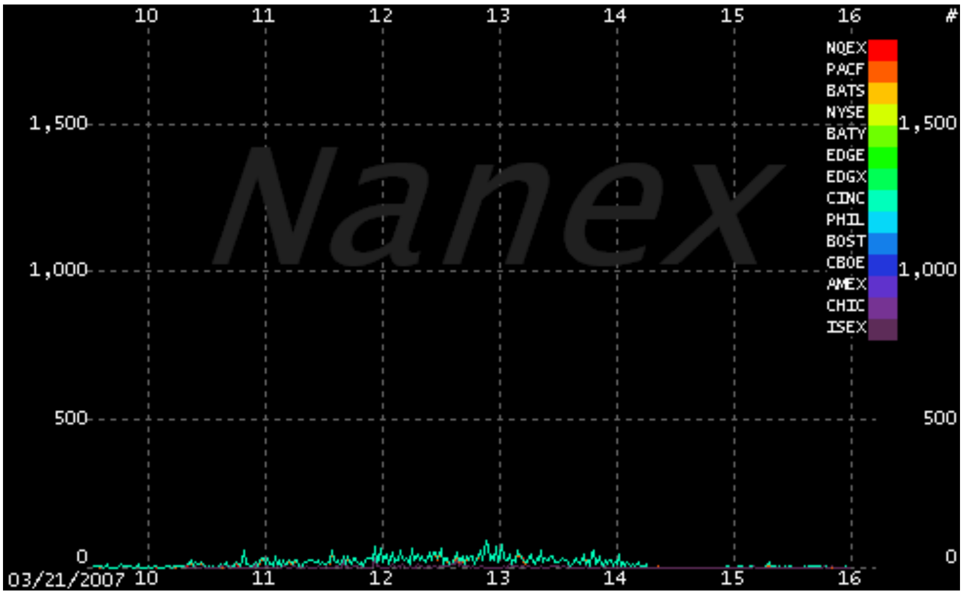

Figure 1: HFT activity for U.S. stock exchanges, as identified by the legend, on March 21, 2007. The x and y axes represent market trading hours—from 9:30 to 16:00 Eastern Time—and number of packets, respectively.

single digit *millisecond* round-trip time, which is the time from data packet transmission to reception [4].

HFT currently generates approximately 35% and 70% of UK and U.S. equity trades, respectively, and thereby exerts significant influence over the global financial markets [5, 6]. In European equity trading, HFT market shares range from 13% to 40%, respectively, for Chi-X and Nasdaq OMX [7]. HFT adaptation rates have been swift and profound. Figures 1 and 2 illustrate high-frequency trading activity for U.S. stock exchanges during March 2007 (top) and January 2012 (bottom), respectively [8].

The ubiquitous diffusion of HFT systems proceeds in contrast to the discord and controversy engendered by those systems. Financial regulators, market participants, academics and members of the media express diverging, and sometimes exaggerated, views with respect to HFT benefits and risks. One point of contention is high-frequency quoting, which involves the rapid oscillation of bid and/or ask prices. HFT proponents consider this technique to be a source of tight bid–ask spreads, which are associated with low volatility [9]. But high-frequency quoting has also been associated with the *May 6, 2010 Flash Crash* [10], a period of extreme volatility [11].

The events of May 6, which temporarily erased $1 trillion in market

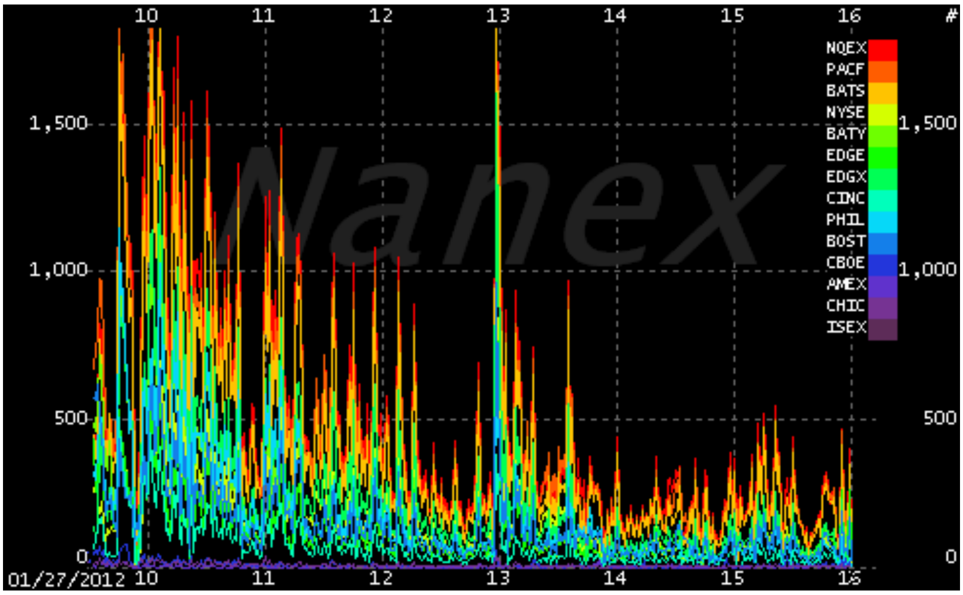

Figure 2: HFT activity for U.S. stock exchanges on January 27, 2012.

value, highlight the potential for amplification of financial anomalies by HFT systems. Other comparable events include the Dow Jones Industrial Average flash crash, which occurred September 29, 2008 [12]; the cocoa futures mini crash, which occurred on March 1, 2011; and the dollar–yen sell off, which occurred on March 16, 2011 [13]. The parameters that characterize these events, including acute price fluctuations that approach upper and lower equity price limits, are clearly attributable to automated processes [14].

We argue for an informed dialogue between stakeholders to realize a common domain of discourse, and an appropriate and resilient regulatory framework. To this end, we have conducted a survey that collates the opinions of academics, market participants and regulators, with respect to HFT benefits and risks. Our findings are presented in the remainder of this report, which is structured as follows. Section 2 presents and evaluates various candidate definitions for HFT. Section 3 describes the ecosystem that underpins HFT activity. Sections 4 and 5 describe the effects of HFT on liquidity and volatility, respectively. The events of May 6 are described in Section 6. Finally, Sections 7 and 8 present recommendations and conclusions, respectively.

2 Defining HFT

A clear and consistent definition for HFT remains elusive [7]. While HFT can be considered a subset of algorithmic trading, HFT and algorithmic trading can also be considered equivalent. The latter viewpoint is justifiable when behavior replaces speed as the primary HFT descriptor. The European Commission (EC) defines algorithmic trading as the use of computer programs to enter trading orders, where the algorithm decides on aspects of order execution such as timing, quantity and price [15]. But HFT has yet to be defined by the EC, and by regulators in general. *High-frequency traders* (HFTs) consider HFT a technique that can be applied to various, low latency, strategies [16]. A technique is the mechanism by which a strategy is executed; a strategy is an algorithm designed to achieve a set of objectives.

On a practical level, HFTs utilize sophisticated technology to interact with markets, analyze market data and execute multiple high-speed trades. Each trade may yield fractional gains; HFT net profits for the fiscal year 2008 in the U.S. were estimated by the TABB Group at \$8–\$20 billion. Most positions, which may be held for as little as a few milliseconds, are closed by the end of the trading day. While HFT is particularly applicable to market making and arbitrage, any automated trading strategy can use HFT techniques to optimize execution and mask trading activities from competitors. In this context, establishing a comprehensive definition for HFT becomes a challenging task.

In addition to practical concerns, HFT also engenders philosophical disagreement. HFT can be considered the evolution of trading, which somewhat parallels the relationship between race cars and horse-drawn carriages. Several arguments support this point of view. HFT is perhaps best described as a *race to zero*, where market participants compete and trade at near-light speeds. This operational requirement imposes capital risk on HFT firms, which must assume significant technological investment to remain competitive. As physical limits are reached however, innovative technology becomes commoditized and speed increases yield diminishing returns. Consequently, market and sentiment analyses become the primary determinants of success [17, 18]. But this is also true for algorithmic, and perhaps discretionary, trading in general. Moreover, unlike arbitrage strategies that can be executed in milliseconds, HFTs that assume a market making role are exposed to market risk. But again, this is true for conventional market participants.

HFT can also be considered a significant extension of discretionary trading. From this viewpoint, HFT is a contrived term that simplifies what is an inherently complex subject. In particular, HFT is performed by automated and adaptive algorithms at near-light speeds. These algorithms operate in the context of a global network that encompasses significant constituent components and human–computer interactions [5]. It is arguable that a stochastic state space of this magnitude cannot be delineated in a succinct manner.

Practical and philosophical concerns have implications for ongoing regulatory efforts. Instead of attempting to define HFT, or to delineate what is an inherently complex state space, regulators should be focused on identifying and restricting harmful behavior. Furthermore, if HFT constitutes the evolution of trading, then any requests for responsibility—social or otherwise—must necessarily be applied to conventional market participants as well as HFTs.

3 The Ecosystem

Cliff and Northrop argue that the global financial markets constitute a complex adaptive ultra-large-scale socio-technical system-of-systems [19]. The concepts that characterize this type of system can be described as follows.

- The adaptive behavior of human and software agents, and the emergent[1] behavior resulting from the multiplicity of interactions between those agents, combine to form a *complex adaptive* system.
- An *ultra-large-scale system* is unprecedented in size with respect to lines of code; number of users; amount of data stored, accessed, manipulated and refined; number of connections and interdependencies among software components; and number of hardware elements [20].
- A *socio-technical system* comprises humans and their computational and physical environments.
- A *system-of-systems* is a complex system that comprises independent, significant and self-contained systems.

[1]Emergence is a phenomenon that describes the appearance of complexity from a multiplicity of relatively simple interactions.

Fundamental differences between conventionally engineered systems and the type of system described above limit scientific and engineering knowledge, thereby constraining our ability to implement effective control structures for the global financial markets [5]. It may therefore be impractical to demand individual accountability in the context of failures that are effectively systemic [19]. A more deliberate strategy would be to develop techniques and tools that can address underlying issues affecting the markets including emergent behavior; heterogeneity via organic growth; normal failure; and scaling.

4 Liquidity

Arbitrage enables HFTs to profit from pricing inefficiencies—expressed as wide bid–offer spreads—that are inherent in a fragmented market structure. In the process, HFT systems reduce those spreads by allocating liquidity—the ability to buy or sell an asset without significantly impacting its price—efficiently across multiple trading venues [5, 13]. As spreads narrow transaction costs are reduced, and prices become more consistent, efficient and transparent.

The issue of liquidity is complicated by HFT activity and strategies; market fragmentation; and the concept of phantom, or transient, liquidity [5]. Phantom liquidity is a consequence of high-frequency quoting, which increases order to trade ratios (i.e., the fraction of submitted orders resulting in a trade) [21]. Due to the inherent speed of high-frequency quoting, phantom liquidity is unavailable to average investors. Liquidity complications force investors to access liquidity either via expensive technology or brokers that will assume the costs of trading [5]. This requirement disadvantages small institutional and private investors who may not be able to afford the technological arms race.

Despite the fact that computer trading has improved liquidity in general, HFTs are accused of withdrawing liquidity during volatile markets [5, 13, 21]. This type of investor behavior can create feedback loops with potential to amplify undesired interactions and outcomes [5, 13]. According to Brogaard, there is no evidence to suggest that HFTs withdraw from markets during periods of volatility [5]. But because HFTs constitute primary liquidity providers, their modus operandi, including limited capital and ultra-fast execution, can engender periodic illiquidity.

4.1 High-Frequency Quoting

HFTs achieve narrow bid–offer spreads by means of high-frequency quoting, a technique that strains pre- and post-trade IT infrastructure—including trading, clearing and settlement systems—with traffic volumes that significantly exceed trade message traffic [8]. Figure 3 illustrates growth in quote message traffic for the Consolidated Quote System (CQS), the official electronic data stream for real-time quote information generated by U.S. stock exchanges. Conversely, Figure 4 illustrates lack of growth in trade message traffic for the Consolidated Tape System (CTS), counterpart to the CQS for real-time trade information.

CQS capacity increases are promptly equalized by growth in quote message traffic [10]; this phenomenon is illustrated by Figure 3 with respect to the CQS, where gaps between groups of lines indicate quote message traffic increases. When capacity limits are tested in this manner, latencies resulting from strained infrastructure could interfere with the efficient operation of the markets; for example, pricing and trade settlement delays could obstruct clearing and settlement systems [13]. This type of systemic risk, which may have contributed to the events of May 6 [22], is discussed further in Section 6.2.

5 Volatility

Economic research provides no direct and unambiguous evidence of causality between HFT and increased volatility, which is asset price variability over time [5]. If HFT contributed to volatility, then HFT diffusion should have increased the intraday-to-overnight volatility ratio; but this correlation is not evident. On the contrary, increased HFT activity has been associated with decreased short-term volatility. According to Chaboud et al., algorithmic trading lowers realized volatility in FX markets; and Brogaard concludes that HFT may have reduced volatility on the Nasdaq [13]. While HFT has also been associated with the events of May 6, which constitute a period of extreme volatility, causality between HFT and the events of May 6 has not been established.

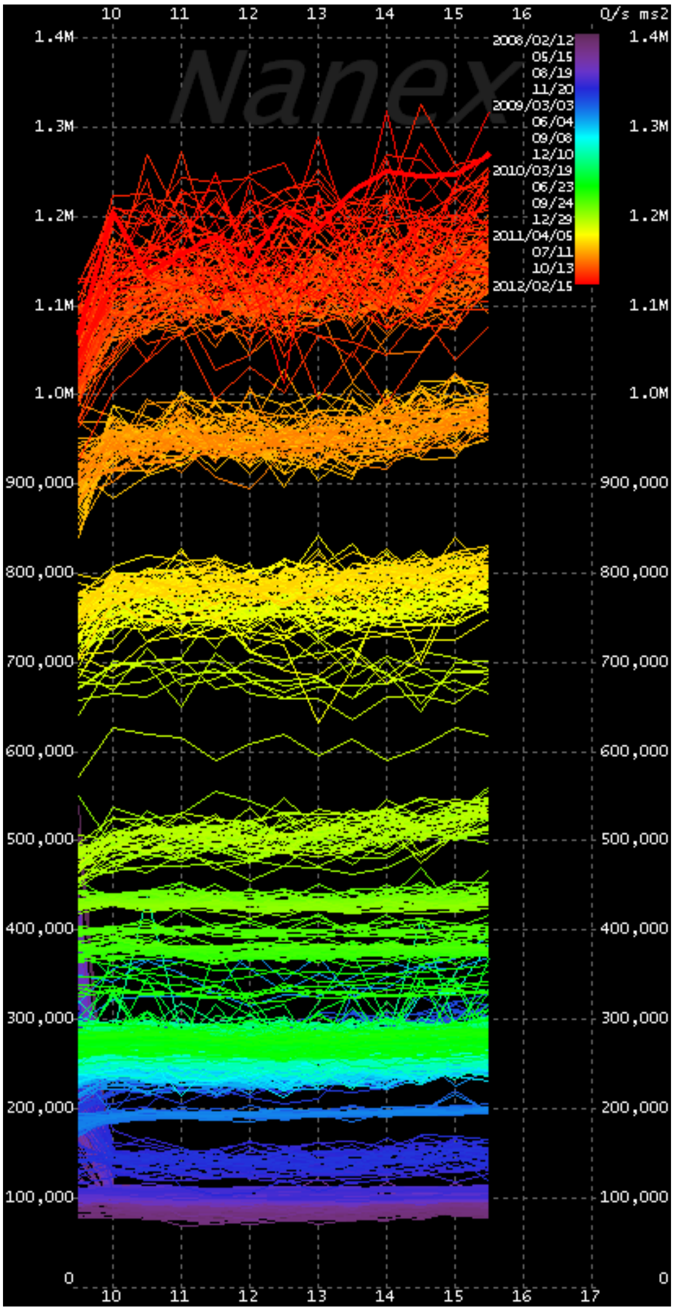

Figure 3: An illustration of CQS two millisecond peak quote rates for the period between February 2008 and February 2012, as identified by the legend. The x and y axes represent market trading hours and quotes per second, respectively.

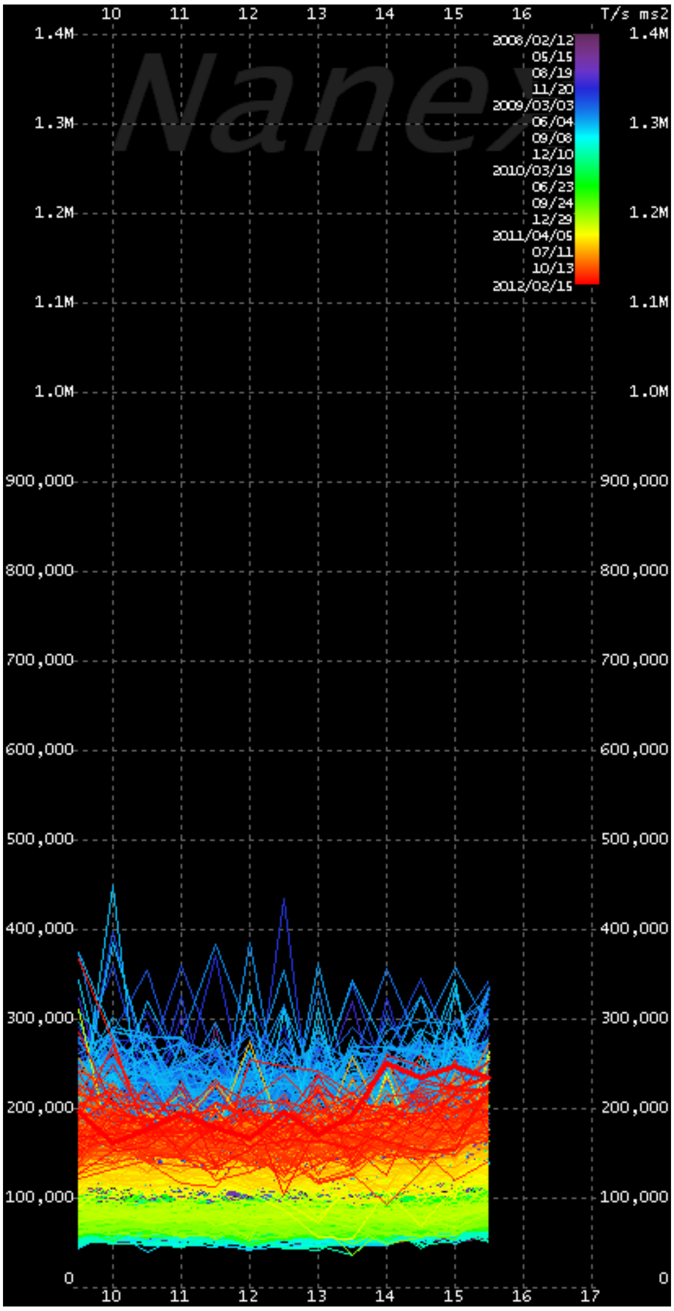

Figure 4: An illustration of CTS two millisecond peak quote rates for the period between February 2008 and February 2012.

6 May 6, 2010

During May 6, major equity indices in futures and securities markets, and almost 8,000 individual equity securities and exchange traded funds, experienced severe price fluctuations [23]. The Dow Jones Industrial Average declined 998.50 points—the biggest intraday point drop in its history—before recovering losses to close down 347.80 points, or 3.2%, at 10520.32 points. Approximately 20,000 trades—across more than 300 securities—were executed at devaluations of over 60%, with many of those trades approaching upper (as high as $100,000) and lower (less than one penny) equity price limits.

6.1 The CFTC–SEC Report

The U.S. Commodity Futures Trading Commission (CFTC) and the U.S. Securities and Exchange Commission (SEC) issued a joint report regarding the events of May 6. These events can be summarized as follows.

Even before the flash crash, May 6 was an unusually turbulent trading day characterized by decreasing liquidity, high volatility and unsettling news regarding the European sovereign debt crisis. At 2:32 p.m., a large *fundamental trader*—a market participant trading to accumulate or reduce a net long/short position—initiated an automated algorithm to sell 75,000 E-Mini S&P 500 (E-Mini) contracts valued at approximately $4.1 billion. The algorithm's execution rate—the rate at which the algorithm submits orders to the market—was set at 9% of the trading volume calculated over the previous minute, disregarding price and time.

The algorithm submitted an initial batch of orders, which was likely absorbed by HFT systems and intermediaries. Between 2:41 p.m. and 2:44 p.m., increased HFT activity, including aggressive selling, prompted the algorithm to increase its own rate of execution. Selling pressure from the algorithm, cross-market arbitrageurs, HFT systems and other traders caused the E-Mini and the S&P 500 SPDR (SPY) to devaluate. Against a backdrop of vanishing liquidity and declining indices, HFT systems generated a *hot potato* volume effect, and acute price fluctuations, by rapidly trading contracts with each other. Following a five second trading pause, which was triggered by the Chicago Mercantile Exchange Stop Logic Functionality in order to prevent further price declines, buy-side interest increased and the

E-Mini and SPY began to recover.

The CFTC–SEC report attributes May 6 to a conjuncture of factors including the automated execution of a large sell order during stressed market conditions; the unreliability of trading volume as an exclusive liquidity indicator; the coupling of derivatives and securities markets; the failure of cross-market circuit breakers; and liquidity erosion resulting from 1) interactions between automated execution programs and algorithmic trading strategies, and 2) the simultaneous withdrawal of multiple market participants.

6.2 The Nanex Theory

Nanex—a financial data vendor and research firm—presents a markedly different point of view [22].

Nine U.S. exchanges—including BATS, CBOE, CHX, ISE, NASDAQ, NASDAQ OMX BX, NSX, NYSE and NYSE Arca—submit bids and/or offers for NYSE listed stocks. The highest bid and lowest offer prices become, respectively, the best available bid and best available offer, collectively known as the National Best Bid and Offer (NBBO) for U.S. equities. Each exchange avoids *crossing* competing exchanges (i.e., bidding higher than other exchanges are offering, or offering lower than other exchanges are bidding) to limit arbitrage opportunities.

At 2:42:46 p.m., NYSE quotes began to queue and thereby fall behind quotes from competing exchanges. This delay caused NYSE bids to cross above the best available offer prices for approximately 250 NYSE listed stocks. HFT systems attempted to exploit this arbitrage opportunity by directing sell order flow to NYSE. While delayed lower priced quotes had yet to be disseminated, sell orders arriving at NYSE were executed against lower priced buy orders from NYSE's designated market maker order book. Because 1) trades and quotes are reported to different data streams (CTS and CQS, respectively), and 2) quote message traffic significantly exceeds trade message traffic (as discussed in Section 4.1), trade data was disseminated with little or no delays. HFT systems detected the sudden price drop and automatically went short, thereby creating an amplifying feedback loop.

Nanex asserts that the events of May 6 were caused by delays in the dissemination of NYSE quotes [22]. However, this causation has *not* been proven, nor is it accepted by the CFTC–SEC Report [23].

6.3 Market Resilience

Reliance on HFT systems to function as primary liquidity providers has positive and negative implications for market resilience. During May 6 and other comparable events, equity prices experienced rapid declines and unprecedented volatility; but losses were reversed by equally rapid and unprecedented recoveries. This type of resilience—the speed with which prices returned to their original levels—was not evident when HFTs were absent from stressed markets; for example, Monday October 19, 1987 [13]. It is also important to note that HFT strategies were largely unaffected by credit constraints resulting from the financial crisis of 2007–2008. Consequently, HFTs remained relatively active, and thereby continued to provide liquidity, during the extended market dislocation of that period.

7 Recommendations

The consensus among regulators and market participants is that HFT activity must not be interrupted. HFTs should be permitted to operate in the markets as long as they do not engage in abusive or parasitic activities that undermine market quality. We advocate that as stakeholders and beneficiaries of a powerful technology, HFTs must assume the responsibility of developing and deploying that technology in an ethical, risk-averse and transparent manner. To mitigate the potential for systemic failures described in Section 3, HFTs should adopt techniques that are specific to the development of high-integrity safety-critical systems. HFTs should also implement transparent systems that can support the regulatory process.

Because regulation compounds the risks that HFTs need to manage, regulators should consider that even minor regulatory adjustments can give rise to significant software development costs and capital risk. Regulation should therefore proceed at a consistent and predictable pace that affords HFTs sufficient time for compliance. To this effect, regulators should engage with market participants, while resisting politically motivated and media induced pressure.

Regulators should also strive to understand electronic trading and HFT by means of empirical data, and not presume all HFT activity to be malicious. While unregulated and untested algorithms are cause for concern, rogue activities are not limited to HFT. The solution to persistent rogue behavior is not the elimination of HFT altogether, but rather the implementation of appropriate monitoring procedures and safeguards. From a technological perspective, market abuse detection systems need not be prohibitively expensive nor disruptive; for example, data tagging—as required by the Markets in Financial Instruments Directive II (MiFID II)—can realize greater transparency by supporting the sophisticated analysis of large and complex data sets.

The importance of data in current HFT environments cannot be overstated [23]. A fair and orderly price discovery process requires robust and timely access to market data for all market participants, whether human or machine. To this end, exchanges must ensure the integrity and reliability of information systems that publish trades and quotes to consolidated market data feeds.

8 Conclusions

The complex technologies and continuous technological advancements that underpin HFT impose significant demands on legislators, regulators and market participants. Regulators must consider specialized, non-financial knowledge in order to establish effective financial regulation. HFTs must consider the severe consequences of irresponsible and harmful behavior on the financial industry, and society in general. But at present, the focus is on impending regulation. If regulators arrive at erroneous conclusions regarding HFT, then regulatory action could damage a technology that has been generally beneficial for the markets.

Our aim has been to review the opinions of academics, market participants and regulators, with respect to HFT benefits and risks. By understanding the risk factors associated with HFT, and the impact of HFT activity on financial stability, stakeholders can realize risk management practices that promote HFT benefits while reducing social costs.

9 Acknowledgements

This report was sponsored by the Association for Financial Markets in Europe (AFME) and the Professional Risk Managers' International Association (PRMIA). The authors would like to thank PRMIA's Board of Directors for their input and support. The authors would also like to thank—in alphabetical order—the following individuals for their contribution.

- Anthony Kirby (Director of Regulatory Reform and Risk Management at Ernst & Young)
- Clive Furness (Managing Director at Contango Markets)
- Evangelos Benos (Research Economist at Bank of England)
- Jasper Jorritsma (Policy Officer at the European Commission)
- John Mueller (Managing Director at Getco Europe)
- Jonathan Howitt (PRMIA Publications Committee)
- Kee-Meng Tan (Managing Director at Knight Capital Europe)
- Mike O'Hara (Co-Founder & CEO at HFT Review)
- Paul Hinton (Partner at Kemp Little LLP)
- Sam Tyfield (Partner at Katten Muchin Rosenman UK LLP)
- Tim Rowe (Manager of Trading Platforms and Settlement Policy at the Financial Services Authority)
- Walter Hendriks (Principal Advisor at Financial Markets Advisory)
- William Rhode (Principal, Director of Fixed Income at TABB Group)

References

[1] I. Aldridge, *High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems.* John Wiley & Sons, Inc., 2010.

[2] London Stock Exchange Group, "London Stock Exchange Group Response to CESR Call for Evidence on Micro-Structural Issues of the European Equity Markets." [online] Available at: `http://bit.ly/14AAHXq`, April 2010.

[3] G. Nuti, M. Mirghaemi, P. Treleaven, and C. Yingsaeree, "Algorithmic Trading," *Computing Now*, 2011.

[4] V. (Willy), "The Value of a Millisecond: Finding the Optimal Speed of a Trading Infrastructure." TABB Group, April 2008.

[5] The Government Office for Science, London, "Foresight: The Future of Computer Trading in Financial Markets." [online] Available at: `http://bit.ly/UvGE4Q`, 2012.

[6] S. Lee, "New World Order: The High Frequency Trading Community and Its Impact on Market Structure." Aite Group, LLC, February 2009.

[7] P. Gomber, B. Arndt, M. Lutat, and T. Uhle, "High-Frequency Trading." [online] Available at: `http://ssrn.com/abstract=1858626`, 2011.

[8] Nanex, "The Rise of the HFT Machines." [online] Available at: `http://www.nanex.net/aqck/2804.html`, 2012.

[9] K.-H. Bae, H. Jang, and K. S. Park, "Traders' Choice Between Limit and Market Orders: Evidence From NYSE Stocks," *Journal of Financial Markets*, vol. 6, pp. 517–538, August 2003.

[10] Nanex, "HFT Is Insatiable." [online] Available at: `http://www.nanex.net/research/cqscap/cqscap.html`, 2011.

[11] A. A. Kirilenko, A. S. Kyle, M. Samadi, and T. Tuzun, "The Flash Crash: The Impact of High Frequency Trading on an Electronic Market." [online] Available at: `http://ssrn.com/abstract=1686004`, May 2011.

[12] Nanex, "The Other Flash Crash." [online] Available at: `http://www.nanex.net/research/29sep2008/29sep2008.html`, 2012.

[13] W. Barker and A. Pomeranets, "The Growth of High-Frequency Trading: Implications for Financial Stability." [online] Available at: `http://bit.ly/VqDBax`, June 2011.

[14] D. Easley, M. M. L. de Prado, and M. O'Hara, "The Microstructure of the 'Flash Crash': Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading," *The Journal of Portfolio Management*, vol. 37, pp. 118–128, November 2010.

[15] D. G. I. Market and E. C. Services, "Review of the Markets in Financial Instruments Directive (MiFID)." [online] Available at: http://bit.ly/gAj3gl, December 2010.

[16] P. Kumar, M. Goldstein, F. Graves, and L. Borucki, "Trading at the Speed of Light: The impact of High-Frequency Trading on Market Performance, Regulatory Oversight, and Securities Litigation." [online] Available at: http://www.brattle.com/_documents/uploadlibrary/upload932.pdf, 2011.

[17] J. Bollen, H. Mao, and X. Zeng, "Twitter Mood Predicts the Stock Market," *Journal of Computational Science*, vol. 2, pp. 1–8, March 2011.

[18] P. A. Hafez and J. Xie, "Factoring Sentiment Risk Into Quant Models." [online] Available at: http://ssrn.com/abstract=2071142, 2012.

[19] D. Cliff and L. Northrop, "The Global Financial Markets: An Ultra-Large-Scale Systems Perspective." [online] Available at: http://bit.ly/GDhfuV, 2010.

[20] P. Feiler, R. P. Gabriel, J. Goodenough, R. Linger, T. Longstaff, R. Kazman, M. Klein, L. Northrop, D. Schmidt, K. Sullivan, and K. Wallnau, *Ultra-Large-Scale Systems: The Software Challenge of the Future.* Carnegie Mellon Software Engineering Institute, 2006.

[21] T. Rijper, W. Sprenkeler, and S. Kip, "High-Frequency Trading." [online] Available at: http://www.optiver.com/corporate/hft.pdf, 2010.

[22] Nanex, "Analysis of the "Flash Crash"." [online] Available at: http://www.nanex.net/20100506/flashcrashanalysis_intro.html, 2010.

[23] U.S. Commodity Futures Trading Commission and U.S. Securities and Exchange Commission, "Findings Regarding the market Events of May 6, 2010." [online] Available at: http://www.sec.gov/news/studies/2010/marketevents-report.pdf, 2010.