



Impersonal efficiency and the dangers of a fully automated securities exchange

The Future of Computer Trading in Financial Markets - Foresight Driver Review – DR 11

Contents

Impersonal efficiency and the dangers of a fully automated securities exchange.....	1
Contents	2
Impersonal efficiency and the dangers of a fully automated securities exchange.....	4
I. Introduction	5
II. History: how automation produced	6
The current market structure.....	6
The computer arrives at Wall Street	6
First efforts at market integration in the United States.....	7
Early challenges to the central position of the London Stock Exchange	8
From bodies to screens in the United Kingdom	9
The rise of electronic trading networks in the US	10
The NYSE scandals and growth of alternative trading venues.....	11
Reg-NMS and the rise of a speed-driven network.....	12
Dark pools and the problems of transparency	13
III. The present.....	13
Risk of a weak-norms market	14
Risk of toxic transparency.....	15
Risk of fragmented innovation	15
Impersonal efficiency and the 2010 Flash Crash.....	15
Weak norms.....	15
Toxic transparency.....	16

Impersonal efficiency and the dangers of a fully automated securities exchange

Fragmented innovation	17
IV. The future	19
Risks stemming from a weak-norms environment	19
Opportunistic trading.....	19
Pre-emptive behaviour.....	20
Overregulation	20
Risks stemming from toxic transparency	21
Increased cognitive mediation	21
Speed and the 'reality gap'	21
Challenges of non-price information	22
Potential risks stemming from fragmented innovation	23
Technical failure and the erosion of institutional confidence.....	23
Risk relief mechanisms	24
Bringing back the social dimension.....	25
Bringing back human judgment	25
Learning from instances of market stress	25
References.....	27

Impersonal efficiency and the dangers of a fully automated securities exchange

Daniel Beunza (London School of Economics), Donald MacKenzie (U. of Edinburgh), Yuval Millo (LSE) and Juan Pablo Pardo-Guerra (LSE)

7 June, 2011

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

I. Introduction

This report identifies impersonal efficiency as a driver of market automation during the past four decades, and speculates about the future problems it might pose. The ideology of impersonal efficiency is rooted in a mistrust of financial intermediaries such as floor brokers and specialists. Impersonal efficiency has guided the development of market automation towards transparency and impersonality, at the expense of human trading floors. The result has been an erosion of the informal norms and human judgment that characterize less anonymous markets. We call impersonal efficiency an ideology because we do not think that impersonal markets are always superior to markets built on social ties. This report traces the historical origins of this ideology, considers the problems it has already created in the recent Flash Crash of 2010, and asks what potential risks it might pose in the future.¹

Before considering its risks, it is important to point first to the many benefits of automation. The most important advantage has been a notable narrowing of the spreads in the equities market. In addition to lower transaction costs, the structure of the market now has competing centres for order matching, and provides direct access to small investors. Equally important, the audit trail generated by electronic trading has made surveillance more effective².

1 The terms automated trading and automated markets are not easily defined, as they capture events in a decades-long history of technology and politics (as this paper describes). However, in spite of this complexity, we aim to define automated trading and markets as financial markets where significant portions of the traded volumes are generated by automatic, computer-based algorithms (high-frequency trading falls under this definition) and/or where significant parts of the order matching and price discovery processes are also conducted using computer-run algorithms.

2 The effects of automation on the behaviour of market order flows is clear. The well-known statistic produced by the Tabb Group, for instance, suggests that up to 77% of the volume of trades in London is originated by high-frequency trading, perhaps the most overt form of automated trading. Broader effects, however, are also discernable. As noted by Robert Barnes of UBS, there is a consistent trend in of increasing traded volumes, which is associated to a progressive reduction in the size of the average trade (see below, from Barnes 2008). The huge increase in the number of bargains can be explained in terms of the increased use of algorithmic trading systems that seek to minimize market impact.

Volumes and bargains at the London Stock Exchange, 1998-2007

Order book average size (£)	Order book number of bargains
2007 14,908	134,150,345
2006 19,362	78,246,367
2005 20,463	51,415,546
2004 21,472	40,771,163
2003 21,739	32,897,427
2002 28,126	23,839,550
2001 41,283	15,750,253
2000 61,749	8,594,471
1999 63,020	5,374,520
1998 58,508	3,583,128

Impersonal efficiency and the dangers of a fully automated securities exchange

Automation, however, has also given rise to potential dangers. By favouring algorithms over human discretion and social cues, the ideology of impersonal efficiency has fundamentally altered the functioning of markets and opened the door to problems. Specifically, the risks of impersonal efficiency manifest themselves in three ways. First, the specific form taken by market automation has undermined the social mechanisms of norm enforcement traditionally deployed by human intermediaries. Second, pursuing transparency by enforcing a system of committed, visible quotes on electronic screens that traders cannot back away from can be self-defeating --in that it can deter from participation in the market. And third, it has promoted innovation in a manner that each actor has a very limited view of the entire landscape. As a result, their decisions can easily have unintended consequences and prompt crises.

Accordingly, we argue that impersonal efficiency gives rise to three forms of risks. First, the risk of a market characterized by weak norms, where the content of norms and their enforcement are uncertain, leading to opportunism and volatility. Second, the risk of toxic transparency: by reducing trading to a mechanism of information exchange, it has undermined the qualitative role of the price system as a coordinator of the market. And third, impersonal efficiency gives rise to the risk of fragmented innovation, that is, situations where partial knowledge about the technological system leads to misinterpreted cues and market uncertainty. This documents expands on this assessment with a history of market automation in the US and UK, an analysis of the Flash Crash and an assessment of the present.

II. History: how automation produced

The current market structure

To understand the present shape of equity markets, this report conducts a brief analysis of their evolution the past four decades. It focuses on two parallel mid-century developments in the UK and the United States. The first of these is the post-war introduction of the computer into the economy and economics (Mirowski 2002). This led to a novel view of markets as information-processing machines rather than social entities. The second development is a resolved mistrust of market intermediaries on the part of regulators. Across both sides of the Atlantic, regulators feared the market power of the central exchanges, and saw their key actors -- the market makers and specialists -- as monopolistic extractors of economic rents (Shapiro 1987). The combined effect of these two trends was the legal imposition of information technology as a form of control.

The computer arrives at Wall Street

Impersonal efficiency and the dangers of a fully automated securities exchange

It did not take long for the post-war computer revolution to arrive at financial markets. Perhaps no text better illustrates the utopian promise of an automated exchange than Fischer Black's 1971 study, titled "Toward a Fully Automated Stock Exchange" (Black 1971; also see Domowitz 1993). Black, who had pioneered the computerization of libraries and hospitals, saw markets as fundamentally a price system that could be reduced to an electronic network of participants linked by a cable, leading to a "thinking whole." This can easily be observed in the figure that was used to illustrate the texts by Fischer Black's article on an automated exchange (see Fig. 1 at the end). In line with his view, Black argued for replacing NYSE's specialists by a central computer.

The information-processing perspective gained an ally in regulatory attempts to improve the market by targeting financial exchanges. These attempts go back to the early 1960s in the United States. At the time, the Securities and Exchange Commission became dissatisfied with the wide spreads quoted by dealers in the over-the-counter market. In its Special Study of the Securities Markets of 1963 (SEC 1963), the SEC concluded that a loose network of dealers connected by telephones and restrained by self-regulation was too difficult to supervise. The SEC thus favoured replacing intermediaries with information technology, requiring electronic quote dissemination at the NASDAQ in 1971 to trace the "memory of the market" for better surveillance (Seligman 2003).

First efforts at market integration in the United States

Regulatory interventions extended to the New York Stock Exchange following the 'Back Office Crisis' of 1968-1969. In this episode, physical operational procedures of clearing and settlement could not keep up with the growth in stock trading volumes (Seligman 1985, Blume et al 1993). Following the crisis, the US Congress asked the SEC to prepare a report examining the causes of the crisis and its implications. In the resulting Institutional Investor Study of 1971 (SEC 1971), the SEC resolved to set up a central market system -- a central arena where all intermediaries could compete under conditions of transparency -- that was not unlike Fischer Black's idea. According to the regulator,

The participation of competing dealers in the central market will ...reduce the element of monopoly power which has accompanied past efforts to establish a central market and will make it possible for potential abuses of such monopoly power to be controlled not only by regulation but to an increasing degree by competition (SEC 1971, pp. xxii-xxiii).

The general motivation for the SEC's initiative was the 'excessive' concentration of power that human market intermediaries accumulated and the biases that this structural characteristic introduced. The suggested solution was a central market where, it was hoped, the sheer number and diversity of intermediaries would prevent any of them from becoming too powerful.

This thinking led to a proposal in 1972 to tie together the various American exchanges in 'a system of communications'. The first substantial policy implementation coming out of this general initiative was the "Consolidated Tape Association plan" that aggregated and recorded trades. This started in 1975, and was run by the Securities Industry Automation Corporation (SEC 1974). Setting up a consolidated tape and a non-exchange entity to run it implied that the prices produced by different exchanges could be compared more effectively than was previously possible.

Impersonal efficiency and the dangers of a fully automated securities exchange

In addition, the SEC introduced in January of 1978 a rule that required the public dissemination of quotations by exchanges and NASD market makers and the calculation of a National Best Bid or Offer (NBBO)(SEC 1978a). This figure was to serve as a benchmark for market transparency of market and, ultimately, for the success of efforts to implement impersonal efficiency. However, the desired transparency could not rely only on unified measures, but on their effective dissemination. Also in January 1978, the SEC called for the ‘the prompt development of a comprehensive market linkage and order routing system to permit the efficient transmission of orders for qualified securities among the various market centers’ (SEC 1978b). This call led to the development of the Inter-market Trading System (ITS) (SEC 1978c), which allowed dealers on one market to transmit orders to another when a better price quote was displayed in the other market.

Gradually, then, all American stock exchanges (and the Chicago Board Options Exchange) joined the national system in a seemingly fluid connected network. Although the system increased transparency in American stock markets, it contributed little to competition, as the NYSE was still able to exploit its pre-eminence and absorb most of the trading volume.

Early challenges to the central position of the London Stock Exchange

A similar picture emerges from the historical developments in the UK. Up to the 1970s, the British financial system was largely dominated by a culture that upheld human intermediaries as necessary for the proper functioning of the stock market. The importance of intermediaries manifested itself in the centrality of the premier intermediary institution, the London Stock Exchange. The LSE was a leading institution among British exchanges³. It was the central trading venue for government debt (gilts), and had forged strong relations with the two key institutions of British finance, the Bank of England and the Treasury. As a mutualised institution, the LSE operated as a trade organisation, deciding through its control over membership who could access the London market. As a physical trading venue, the LSE provided the marketplace itself in the form of a trading floor restricted to members and a handful of information providers. As a clearinghouse, the LSE settled the deals struck on its floor. As a listing authority, the LSE was the obligatory passage point for company news -- as well as for issuing shares and raising capital in the market. And, through its mechanisms of surveillance, the Exchange worked as the regulator, ensuring confidence in, and the stability of, the market. In summary, the central position of the London Stock Exchange meant that personal judgment, centralization and self-regulation were the keystones of the market (Michie 2001, Pardo-Guerra 2010).

Starting in the 1970s, however, the UK system also exhibited a drive towards impersonal efficiency. Challenges to the LSE’s role as a de facto essential intermediary emerged as early as 1973, when merchant banks opposed the fixed commissions charged by member brokers of the LSE. This led to the development of an independent electronic trading platform based on the American Instinet. The system, known as Automated Real-Time Investments Exchange Limited (ARIEL), did not capture considerable volumes. But it led to a reduction of commissions for institutional users.

³ Since the amalgamation of British stock exchange in 1973 and the subsequent creation of the Stock Exchange of Great Britain and Ireland in 1973, the LSE was in to all practical extent synonymous with financial activity in the United Kingdom. See Thomas (1973).

Impersonal efficiency and the dangers of a fully automated securities exchange

In the late 1970s, a related pressure to disintermediate took the form of anti-trust concerns about the market power of the LSE. In 1978, the Office of Fair Trading ruled that the LSE's rulebook and code of practice infringed competition laws. These regulatory pressures resulted in a decision to overhaul the LSE's structure and operations in 1986, in what is often referred to as Big Bang. The Big Bang entailed the dismantling of fixed commissions on brokerage, eliminating single capacity (whereby market-making and broking were segregated), allowing non-nationals into the membership, and introducing an electronic dealing system.

From bodies to screens in the United Kingdom

Another important step towards impersonal efficiency in the UK was the development of SEAQ, a system that was meant to emulate the operation of the trading floor through a network of screens and phones, to "fold" into silicon and code a system of flesh and blood (Pardo-Guerra 2011). Given that SEAQ was a quote-driven system based on negotiations over the telephone, however, the LSE remained committed to human intermediation as a form of market stabilization: even the small order automated execution system developed by the Stock Exchange in 1989 contemplated the option of diverting orders to a series of nominated counterparties, so keeping the business relationships of the past in an entirely automated environment (according to some market participants active at the time, this feature made the LSE's automated execution system more attractive than other competing platforms).

Unlike the United States, where the financial system was split between several venues, trading in British equities was largely contained in the LSE. The regulatory pressure for increased competition that emerged in the UK in the mid-1970s applied mostly to activities within the LSE and had not provided incentives for the emergence of direct competitors to the organization. In the mid-1990s, the importance of the LSE and the human intermediary was reduced again with a transformation of market microstructure that introduced competition in the market. First, there was a transformation of the regulatory environment in Britain that included the adoption of economic theory as source of analytical inputs⁴ and the acknowledgement among leading regulators that the City was lagging behind other financial centres and required a signal of innovation.⁵

A second source of regulatory pressure was the development of the first alternative to the LSE, Tradepoint, which introduced the electronic order book to the British financial community. It also accelerated the LSE's migration from intermediary-driven SEAQ system to a price-driven system called SETS in 1997. It is important to understand the manner in which Tradepoint

4 Notably, from the late 1980s, the Office of Fair Trading and other government bodies sought increasing advice from economists. The groups at City University and the Financial Markets Group of the London School of Economics were particularly active in conducting studies of the stock market which reverberated with the views of the OFT. To a considerable extent, policies forged on and after 1990 were increasingly informed by a particular neoclassical conceptualization of transparency as necessary for maintaining fairness in the market, one which harmonized with general notion of impersonal efficiency.

5 Such impact of economic theory on 'real world' issues such as market design is addressed in the sociology of finance literature through the concept of performativity of economics (for example, MacKenzie and Millo 2003, MacKenzie 2006, Millo and MacKenzie 2009). The performativity concept acknowledges that economic theories, when incorporated into rules, procedures or technological artifacts may affect and shape the behaviour of markets and not only analyze it.

embodied the principles of impersonal efficiency⁶. Introduced in 1995, Tradepoint was a brainchild of several former Stock Exchange technologists who, after leaving the LSE in the early 1990s, converted their expertise into a concrete commercial product. For these technologists, human intermediation was largely deemed unnecessary. Indeed, as early as 1980, some of the individuals who were involved with the development of Tradepoint had envisaged moving the Stock Exchange to an all-electronic order-driven and globally-accessible computer network. The network, represented in figure 2, embodies the technological aspects of impersonal efficiency: access, speed and decentralization. As an electronic platform for exchange, Tradepoint replicated the views on the market held by these technologists.

Tradepoint introduced the first version of anonymous direct market access to London. In this sense, it meant to provide a trustworthy, yet impersonal, mechanism of exchange for buy-side institutions. For regulators, Tradepoint was a symbol of innovation: although it could have operated as a trading platform, the founders were encouraged by key members of the civil service to register the system as a Recognized Investment Exchange.

Tradepoint's use of an electronic order book demonstrated the possibility for a new form of trading. Instead of intermediation over the phone, the electronic order book entailed direct access to the matching engine. Indeed, two firms that today are quite active in the domain of high-frequency trading (Timber Hill and Optiver) were among the first to 'plug into' Tradepoint's engines as early as 1995. Along with broader changes in the organizational environment in the British financial services industry, the move to an order-driven market triggered increased competition, reduced spreads in the most active shares, and consequently lower profits per trade, three factors that define the landscape of fragmented innovation of today's markets.

The rise of electronic trading networks in the US

Two events bolstered the initial moves towards an integrated network in the United States during the 1990s. In the wake of the 1987 market crash, the market makers at the NASDAQ came to be seen as opportunistic because they did not fulfil their obligations to supply liquidity and ensure continuous trading. (By contrast, at least some NYSE specialists appear to have tried to honour their commitments) (US Treasury 1988). Thus, although the origin of the 1987 market crash was attributed at the time to "program trading,"⁷ there was also a reading of the event that reinforced regulatory mistrust of market makers (Greenwald and Stein 1988). It is important to clarify that program trading is not exactly the same as algorithmic trading. There is a sense indeed in program trading was partly based on a portfolio insurance algorithm. However, the orders were not executed in an automated way. The premier portfolio insurance firm, Leland, O'Brien and Rubinstein, would decide on its trades on the basis of an algorithm, but then they would contact a broker in Chicago, who would call the floor and execute the order. In short, the algorithms were placed at the very start of a long chain of human interaction.

Scepticism about human intermediaries view was greatly reinforced in 1994 with the publication of Christie and Schultz's article on the NASDAQ (Christie & Schultz 1994). The

⁶ Interviews with three Tradepoint developers; London June 2007; London January 2011 and London February 2011.

⁷ The authors of this document do not agree with that.

Impersonal efficiency and the dangers of a fully automated securities exchange

finding that market makers colluded in avoiding quoting in odd eighths⁸ led to an investigation and the eventual adoption of the SEC's Regulation ATS (Alternative Trading Systems) in 1996. It recognised the existence of "Electronic Communication Networks" and gave them direct access to the inter-exchange order-routing system -- and, as a result, to a much wider audience of potential investors than before. Part of regulation ATS was the "Limit Order Display Rule" that required exchange specialists to display publicly customer limit orders that improve on the specialist's displayed price or size (SEC 1996). The intention behind the rule was to increase transparency among the different markets that supplied quotes to the communication system.

The concurrent development of trading networks during the 1990s shows that impersonal efficiency is not simply a regulatory driver, but also a force that shaped the strategies of private companies. For instance, circa 1984 Goldman Sachs assembled a team of microstructure specialists, including the electronic market visionary Fischer Black. These began developing electronic trading networks under the direction of Hank Paulson and Duncan Niederauer, and the latter would eventually leave his imprint on the system by heading the NYSE.

Nevertheless, the growth of trading networks remained limited for as long as a manual exchange such as the NYSE remained the central piece of the system. The trading networks argued that their systems could not be programmed so as to wait for manual execution in other markets. This motivated the SEC to develop a new set of rules that would allow the electronic markets more prominence in the central market system.

The NYSE scandals and growth of alternative trading venues

The agenda of the SEC gained political momentum in the wake of the "Grasso scandal" at the NYSE. On August 2003 it was reported that the board of the NYSE had agreed to grant the CEO, Richard Grasso, a compensation package of \$139 million. The revelation of the sheer amount --larger than the combined CEO compensation of the ten largest competitors of the NYSE -- prompted the suspicion that the board suffered a conflict of interests. Moreover, the credibility of the NYSE's role as a critical gatekeeper between firms and the investing public was damaged as a result of the accusations. Following a continuous public pressure, the board decided to ask Grasso to resign (Gasparino 2007). Grasso's resignation did not signal the end of the affair. The SEC, along with New York State's attorney general office started an investigation that led to the submission of a civil law suit against Grasso in May 2004, accusing him and other board members of manipulating governance procedures at the NYSE board.⁹

The legitimacy of the NYSE suffered another setback with the so-called "specialists scandals." Following an internal investigation, in October 2003 the NYSE fined five of seven specialist firms \$150 million for habitual abuse of their market roles. The common abuses among the specialists were of 'inter-positioning' (placing an order at a price between current bids and offers) or of 'front-running' (upon receiving, trading for the specialist's account and only then

⁸ The research found that the market makers tended to 'ignore' the possibility to quote prices that ended in odd eighths of a dollar (the tick size allowed at the period) and quoted only prices that ended in 'even' eighths (2/8, 4/8, 6/8).

⁹ The lawsuit against Grasso was discussed in the New York Supreme Court, which ordered Grasso in October 2006 to repay to the NYSE part of the compensation package. This decision, however, was reversed in 2008 in an appeal by Grasso.

Impersonal efficiency and the dangers of a fully automated securities exchange

executing the order). In addition to the fines, the specialists firms agreed to pay \$240 million as part of a settlement with the Exchange. Following these events, the NYSE itself was accused of failing to regulate the specialists and the SEC started an enforcement action against the exchange, which resulted in a settlement that required the NYSE, among other things, to install surveillance devices and fund a regulatory auditor at the exchange (Ellis et al 2009). These events, which followed closely the Grasso scandal, contributed to weaken the standing of the Exchange as a self-regulating body.

Following the departure of Grasso, the NYSE appointed John Reed as interim Chairman, and eventually, in 2004, nominated John Thain, then President of Goldman Sachs and a longtime proponent of electronic trading, as Chairman. Following several reforms, the NYSE shifted its strategy from the trading floor to electronic trading. These included demutualization, the acquisition of Archipelago Exchange, and the merger with Euronext. The shift was completed with the appointment of Duncan Niederauer, also from Goldman Sachs, who concentrated the capital investments of the NYSE in two new data centers that reduced latency and allowed for co-location (Lucchetti 2007).

Reg-NMS and the rise of a speed-driven network

Amidst these developments at the NYSE, the SEC launched a major initiative that would radically overhaul trading in the US around the principles of speed, transparency and network interconnectedness. This set of rules, known as Regulation National Market System (Reg-NMS) was presented initially in 2004 and was enacted in 2005. The initiative sought to introduce competition by improving the interconnectedness among American exchanges. One of the key aspect of Reg-NMS was a clause known commonly as the 'trade-through rule', which demanded full disclosure and quick tradability of all prices existing at the different markets that were part of the system. This meant that 'fast' markets – ones where orders were executed electronically were to be connected with 'slow' markets, where humans processed orders. However, allowing (human) traders in the slow markets the chance to react to orders from the fast markets was regarded as giving them an unfair advantage. According to Lawrence Harris, the Chief Economist of the SEC at the time and the leading force behind Regulation NMS, the situation amounted to granting the human traders at the NYSE a free 'look back option'. In contrast, a national market system where all participating exchanges are given similar reaction time to orders would be fairer and, following this principle, the 'trade through rule' set a reaction time of one second for all broker-dealers.

The direct implication of the fast trading requirement for the NYSE specialists was that they had to transfer most of their operation to automatic trading. In the following years, the market share of the floor of the New York Stock Exchange shrunk dramatically, the number of trading rooms shrunk from five to two, and the number of employees fell accordingly.

Compounding the effect of Reg-NMS was the rise of automated order execution. By slicing an order into small components, a market actor, in effect, diminishes the social component of the trade. Just as a tray of sliced sandwiches removes the need on the part of the participants to an office lunch to compete for a given flavour or engage in barter with their allocated sandwiches (in short, to participate in social engagement), slicing the orders of a market actor removes the need for human intermediaries such as NYSE specialists.¹⁰ Following the

¹⁰ As a strategy to reduce market impact, slicing may be said to create conditions that, under some conditions, make real-world markets closer to their economic ideal types: by reducing the size of the average transaction, slicing effectively reduces the information content of each trade.

implementation of Reg-NMS, the NYSE's ability to affect prices through human intermediation was virtually decimated.

Dark pools and the problems of transparency

The downsides of liquidity became first visible with the development of so-called "dark pools," electronic order-matching platforms that deliberately avoid publicizing bids and asks on offer and delay publication of volumes and prices¹¹. Dark pools actively monitor and control the population of traders, aiming to sift out ones who use order-slicing algorithms on a regular basis (Mittal 2008). Dark pools themselves are regarded by exchanges as a factor that contributes to the deteriorating quality of order execution and price reliability. It is difficult to assess the volumes and other operating data related to dark pools because of their secretive nature and also because of ambiguities in their definitions. However, it can be established that since 2008 there was a sharp rise in the number of dark pools in operation. Among the leading ones are the ones operated by Liquidnet, The Investment Technology Group (which operates POSIT) and Chi-X.

III. The present

In sum, our historical analysis points to a core driver that shaped the evolution of equities markets in the UK and US. This driver is the belief in impersonal efficiency and the inherent superiority of automated trading over human-based trading. This ideology, espoused by regulators and shared by technologists, was a motivating factor behind the elimination of human market intermediaries and the introduction of direct competition among exchanges. These steps, according to the belief, would improve market efficiency and benefit investors. Impersonal efficiency is informed by a central metaphor -- the market as an information-processing network-- that displaced previous perspectives such as the market as a social entity, located in a specific a geographical place. Impersonal efficiency has been supported by the development of new information technology, and guided automation in a way that anonymized actors, atomized orders and privileged speed.

As a driver of market evolution, impersonal efficiency has one key implication, namely, that automation has taken only one path out of many possible alternatives, and it is a path that rendered the market less social. The informants that Beunza encountered in his fieldwork at the New York Stock Exchange in 2009¹² repeatedly insisted in this aspect, and warned about possible dangers of the state of the US equities market. Specifically, they warned against the fragmentation of liquidity, the demise of the obligations of the specialists, the disappearance of

¹¹ Although new as opaque trading platforms, dark pools are not entirely new to the market. To some extent, they follow a logic of internalization of large trades that has existed for some time in the financial system. In the LSE, and despite the division between brokers and market-makers, brokers could match orders if necessary, paying jobbers a commission that represented a fraction of the touch or spread. The process, known as put-through, allowed brokers to conduct relatively large deals whilst minimizing information leakages. Similarly, relatively opaque Inter-Dealer Broker networks have existed in certain global markets for more than half a century and may have served as inspirations for today's dark pools.

¹² This study is informed by research conducted by Daniel Beunza in New York in 2009, as well as a series of interviews conducted by Dr. Yuval Millo with market participants in the US and Dr. Juan Pablo Pardo-Guerra with market participants in the UK.

large blocks, and the dangers of giving up judgment to rely entirely on algorithmic trading and order-matching.

Our own analysis points to three areas of risk in a market characterized by impersonal efficiency.

Risk of a weak-norms market

First, the application of impersonal efficiency may bring about a weakening of market norms. This includes the lack of specialists at the NYSE, but also the danger stemming from anonymous trading and the reduced scope of authority for exchanges to shape their order matching procedures so that it improves market order. We refer to this as the risk of a weak-norms market.

The problems posed by a weak-norms market have been already discussed in the literature. The economic historian Paul David, building on the work of Lawrence Lessig (Lessig, 1999) argues that the root of the problems lies in the replacement of a human network --governed by the principles of community and shared norms-- by an electronic network that is governed by rigid technical codes (David 2010). Indeed, the problem is more general: innovators are frequently not bound by the obligations of the traditional system they aim to innovate. As new practices and devices replace the old ones, the governance of the system is influenced by the innovators who program the technical specifications of the system in code. However, because the innovators tend to have better control than others over the technical knowledge, there is little opportunity for public scrutiny.

This risk is consistent with a growing literature in the sociology of financial markets that argues that shared informal norms (as opposed to formal legal rules) are a basic component of any smooth running market. Baker's seminal work (1984) demonstrates that the size of the crowds on the trading floors have a key role in shaping trading patterns, as it determines the extent to which floor traders can informally enforce norms of orderly market behaviour. Similarly, Zaloom's (2001) study of the introduction of electronic trading at the Chicago Board of Trade concludes that the change in environment from face-to-face trading to screen-based trading resulted in losing a nuance-rich communicative environment. In a pioneering sociological study, Abolafia (1996: 119) argues that NYSE specialists, especially under the chairmanship of John Phelan between 1984 and 1990, developed a "culture of restraint" that led members to follow the norms even when it was against their own self-interest. By contrast, market makers at the NASDAQ lacked the clear "positive obligations" that NYSE specialists confronted. As Ingrebretsen (2002: 98) points out "some market makers on the NASDAQ did as their counterparts had done decades earlier. They simply refused to answer the phone." The implication from these various studies of options and equity traders is that norms matter in the smooth running of markets.¹³

¹³ This perspective is different but not necessarily inconsistent supported by analysis by some behavioral finance scholars who study the relations between social connections and market behaviour. Hong et al. (2005) and Cohen et al. (2008, 2010) examine the trading behavior of professional money managers and find that behavior co-varies more positively when managers are (i) located in the same city and (ii) went to college together, respectively. Similarly, Stein (2008) posits a formal model of bilateral conversations in which actors honestly exchange ideas with their competitors when the flow of information is bi-directional between each pair of actors and when the actors can develop useful ideas on the information shared.

Impersonal efficiency and the dangers of a fully automated securities exchange

Risk of toxic transparency

Second, a market designed around the principles of impersonal efficiency runs the risk of discouraging participation and thus limiting liquidity. Stephen Wunsch has labelled this mechanism “toxic transparency” (Wunsch 2011). In Wunsch’s view, forcing trading orders into electronic screens increases transparency, but does so at the cost of discouraging participation. There are two mechanisms that contribute to this outcome: first, whenever traders make their orders public, it inevitably means that prices will move against them. And, second, forcing traders to make commitments that they cannot back out of in a panic means that traders will not make commitments at all. As Wunsch writes,

It just may be inherent in the nature of transparent electronic screens that liquidity will disappear more quickly from them when traders get nervous than it would have from traditional manual markets. It is the committed, visible, no-backing-away nature of participation on electronic screens that makes participation dangerous (Wunsch 2011).

Thus, in other words, the shift towards electronic screens and transparency runs the risk of being self-defeating. This risk is clear, for instance, in the rise of dark pools following Reg-NMS.

Risk of fragmented innovation

Third, by transferring markets to the electronic medium, the competitive field narrowed down, driving innovation to compete primarily on achieving higher and higher execution speeds. This can lead to an increased importance of secrecy and to the emergence of a market comprised of different actors using different platforms with different technical requirements. The risk in such market is that no-one would know the technical state of the system as a whole, opening the door to actions that have negative unintended consequences. We refer to this as the risk of fragmented innovation.

Impersonal efficiency and the 2010 Flash Crash

Our three categories of risk explain some key aspects of the American Flash Crash of 2010. The following paragraphs summarize the official account of the crash by the SEC/CFTC, as well as competing analyses, showing how the three categories of risks outlined above offer a useful device to understand the events.

Weak norms

One important aspect of the Flash Crash was the decision by several market participants to pull out of the market after the sharp fall in prices at 2.45 pm on May 6th. These actors include high-frequency traders, but also other categories. However, high-frequency traders had their own specific reasons for pulling out. The report (p. 4-5) points out to some of these reasons. These include a fear of erroneous data, impact of price moves on risk and position limits, impact on intraday profit and loss, fear of broken trades, and fear of a cataclysmic event. In a much-cited analysis of the Flash Crash, Stephen Wunsch argues that the crash is partly a result of the lack of discretion within exchanges imposed by Reg-NMS: whereas exchanges had an incentive to protect the integrity of the process before Reg-NMS, the passage of the law eliminated this positive effect. “With reputations on the line,” Wunsch argues, “traders and exchange officials applied discretion based on a code of conduct that vetted each stage of a

trade for reasonability. After reforms were enacted (...) reputations were irrelevant.” By altering the incentives of exchanges, Reg-NMS has contributed to a weak-norms environment.¹⁴

Another reading of the Flash Crash is as an instance of lack of forethought on the part of a single market participant. So intense was the selling pressure of Waddell’s Sell Algorithm that the slicing of the order by the execution algorithm failed in the goal of reducing the price impact. The episode is a reminder that execution algorithms do not eliminate the need for participants to observe social norms of orderly behaviour. Even if execution algorithms reduce the need for human specialists who conceal social cues, there is still a need for norms of orderly behaviour. The implication is that there may still be a need for some norm enforcer of last resort. Or alternatively, for some emergent mechanism of community-driven norm enforcement.

Another example of how legal trades can upset market order is given by Citigroup’s 2004 trade in the European bonds. On the morning of August 2, 2004 Citigroup sought to exploit a discrepancy resulting from differences in the liquidity of the European bond and bond futures markets. Market liquidity, however, was the result of a conscious ‘liquidity pact’ between the banks and Continental European governments using the MTS bond-trading system. The traders used new software to place several orders simultaneously, and concerned that it had not worked in the first place, used it twice. Citigroup made a profit of almost £10 million in the trade, but the bank was widely condemned and the UK Financial Services Authority (FSA) forced it to relinquish the profits from the trade and to pay a further penalty of £4 million. The FSA did not accuse Citigroup of having broken the law, but nevertheless held that its trading had violated two of the authority’s ‘Principles for Business.’

Toxic transparency

The problem of price quality manifested itself during the Flash Crash in several ways. One manifestation can be seen in the “hot potato” effect. A sequence of rapid trades that added to market volume but did very little to absorb sales. “Between 2:45:13 and 2:45:27, HFTs traded over 27,000 contracts, which accounted for about 49 percent of the total trading volume, while buying only about 200 additional contracts net” (p. 2). The report attributes the drop in the liquidity of the market to this high volume:

At this time, buy-side market depth in the E-Mini fell to about \$58 million, less than 1% of its depth from that morning’s level. As liquidity vanished, the price of the E-Mini dropped by an additional 1.7% in just these 15 seconds, to reach its intraday low of 1056. This sudden decline in both price and liquidity may be symptomatic of the notion that prices were moving so fast, fundamental buyers and cross-market arbitrageurs were either unable or unwilling to supply enough buy-side liquidity (p. 7-8).

¹⁴ Qualitatively similar views were presented by market participants elsewhere. In particular, a certain school of thought deems algorithms as inherently incomplete when compared with individual human judgment. As one trader noted, “the behaviour of human traders in general is too complex and diverse to be effectively replicated by an HFT algorithm, and only basic elements and scenarios are coded. There could be a lot of overlap among HFT algos, and where the same ideas have been cloned over many computers, and the diversity of opinion and behaviour could have been lost. Speed has therefore become the main competitive advantage among high-frequency traders. Markets also run a higher risk of liquidity holes. Interview with Dr Karim Taleb, Robust Methods, The High-frequency Trading Review.

Thus, in other words, the report posits a direct negative relationship between speed, volume and liquidity. This is interesting because this relationship has traditionally been considered positive rather than negative: the greater the activity/ volume, the greater the liquidity.

Fragmented innovation

The first lesson from the Flash Crash derived by the authors of the Commission is that there are interdependencies that have previously been overlooked. Chief among these is the “interaction between automated execution programs and algorithmic trading strategies,” which can erode liquidity and lead to a disorderly market (p. 6).¹⁵

Post-hoc, acknowledging the presence of interdependencies is appropriate -- but even if the organization of the market is altered to account for them, the Flash Crash still raises the question of which other interdependencies that have not yet been explored. Furthermore, the notion of fragmented innovation posits that interdependencies are difficult to locate. As a result, structural reforms after one crisis will never be a complete remedy for the next crisis. Another corollary is that situations of crisis or semi-crisis (market stress) are desirable in that they allow market actors to discover interdependencies. At the same time, these crises carry their own cost in terms of legitimacy, so there is a trade off between the two.

Indeed, fragmented innovation in the financial landscape implies that regulators and market participants must be particularly aware of the presence of so-called ‘epistemic accidents’, which are ‘accidents caused by engineering beliefs that prove to be erroneous, even though those beliefs are logical and well-founded’ (Downer 2010a). Unlike normal accidents — which are often associated to tightly-coupled systems, are entirely tractable after the incident and seldom reoccur (Perrow 1999) — epistemic accidents arise from the fact that it is impossible, in principle, to entirely understand a technological system, particularly in highly innovative domains. Thus, epistemic accidents are likely to reoccur after an incident and are less tractable than normal accidents.

There are reasons to speculate that the flash crash resembles such form of epistemic accidents. Along with the timescales involved in reconstructing the event (only to face a lukewarm reception to the Final Report), if the claims by software developer Nanex LLC are correct, in the sense that apparently extreme events were experienced prior to 6 May 2010 (Nanex, 2011), there would be some evidence to suggest that extreme technologically-mediated liquidity events are too common to be deemed normal accidents and may well reflect a deeper shared uncertainty about the operation of the financial system.

¹⁵ Indeed, systemic and widely entrenched forms of uncertainty may well be at the core of current market dynamics, as echoed in the trade literature by several market participants: “people hit the send button without really knowing where things are going. That’s not right and it’s not consistent with the best-ex responsibilities of brokers. So I believe that we’d be in a better world if there were less opaqueness and more transparency on the broker routing front. Now it’s another question whether that needs to be driven by regulation or whether the buy-side just has to speak up more and say, “damn it if I’m using your algorithm you’ve got to tell me where everything is going and I want real time executions by venue and I want full transparency, and I don’t want to hear you tell me you can’t do it because you have a non-disclosure agreement in place with someone”. People need to know where their orders are going and it would benefit the overall market structure and the end investors if there were more transparency there. Interview with Joe Gawronski, Rosenblatt Securities, The High-frequency Trading Review.

Impersonal efficiency and the dangers of a fully automated securities exchange

Here, the key regulatory insight provided by the sociological literature is that one must deploy mechanisms that learn from incidents that are short of accidents and that reveal some of the limitations of the state-of-the-art knowledge of the financial system. Such mechanisms may well be lacking at present. In 2005, for instance, there was an episode called the correlation crisis in the market for CDOs which could have been an opportunity to learn about the systemic behaviour of these instruments. Aside from the odd trade magazine article, there was no investigation or institutional form of learning from this episode.

Given the above, market pauses and circuit breakers can be seen in a different light. There is an important difference between the pauses that individual firms decide to put in place and the market-wide pauses that exchanges decide to put in place. While the individual pause disrupts the market by removing liquidity from the system, the market pause (or the individual stock circuit breaker) allows participants to engage in sensemaking and overcome the structural ignorance. It is key, then, to better understand the precise mechanisms whereby market participants restore a measure of shared understanding when the market pause is in effect. What exactly happens? The question is complex because different participants (i.e., high-frequency traders, institutional investors, arbitrageurs, retail traders) operate in different ways. Indeed, answering the question may require rethinking markets (and regulatory actions) in terms of an intricate combination of differential computational abilities, risk profiles, and network structures. The success of NYSE's liquidity replenishment points may well have derived from their capacity to return the network to normal operating conditions as much as their role as moments to readjust positions.

Another problem highlighted in the report is the greater complexity of aggregating different sources of data in a world of automated execution and trading algorithms. As the report explains, in an automated market,

[Trading] is further complicated by the many sources of data that must be aggregated in order to form a complete picture of the markets upon which decisions to trade can be based. Varied data conventions, differing methods of communication, the sheer volume of quotes, orders, and trades produced each second, and even inherent time lags based on the laws of physics add yet more complexity (p. 7-8).

In other words, there is a high measure of heterogeneity in the format, rhythm and origin of the data. Furthermore, this heterogeneity is not entirely understood, leading to uncertainty. Indeed, the Commission's report argues that although market delays were not "the primary factor" (p. 8) in the events of May 6, delays in market data did pose problems. "Our analysis of that day," the report writes, "reveal the extent to which the actions of market participants can be influenced by uncertainty about, or delays in, market data."

The report points to a link in the system that accomplishes an integrating role, namely, the "market centres" (which we take to mean exchanges) and their data processes "that involve the publication of trades and quotes to the consolidated market data feeds" (p. 8).

The above may well be a veiled reference to the allegation made by certain market participants that the flash crash was the result of unsynchronized data feeds of the New York Stock Exchange (a point, for instance, raised by Nanex). Such delays result from the heterogeneous configuration of the network, itself consequence of fragmented innovation and a feature of structural ignorance. While the SEC's investigations did not find a connection between the events of 6 May and the lags in the NYSE's feeds, the possibility remains that in a highly interconnected and automated system, adverse network conditions may lead to aggressive reactions that are the result of what amounts to a technologically-induced form of mispricing.

Indeed, it is crucial to note that in the case of the American market, any significant lag – where significant can be a question of a few milliseconds – renders meaningless regulatory constructs such as the National Best Bid-Offer. While European markets may be insulated from such events the fact remains that lags and subsequent actions based on mispricing are sources of potential controversy, particularly in moments of high volatility.

Further, as an issue of structural ignorance, it is important to bear in mind that the inability of market participants to recognize the presence of lags in data feeds during a particular market event is contingent on the specific time-stamping policies followed by regulators and trading venues. In their initial report on the Flash Crash, for instance, Nanex noted that the NYSE's problems seemed to stem from the stamping policies of the Consolidated Quotation System, which stamp data on dissemination rather than generation. In effect, just as synchronization is central to maintaining system integrity, proper stamping policies are critical to maintaining systemic confidence. The technological roadmap of German exchange Eurex, as an illustration, is strongly based on making guaranteeing multiple levels of stamping (including pre before and after quotes are fed to both the gateways and the matching engine) and keeping their policies concerning the feeds as clear as possible.¹⁶

IV. The future

Looking ahead, in ten years we expect to see an evolved form of the risks outlined above if the present trend towards impersonal efficiency continues. Weak norms, toxic transparency and fragmented innovation will worsen and manifest themselves more acutely. We detail these below.

Risks stemming from a weak-norms environment

As systems integration grows over time, decision-making will increasingly distribute across actors. As a result, responsibility will also become more and more distributed. As it gets harder and harder to attribute blame to any single actor, market participants will find it easier to disregard market norms.

Opportunistic trading

A situation of weak norms may lead to deviant behaviour on the part of market participants. Deviant behaviour simply calls for more deviance. In sociology, the effect is known as the “broken window” theory: once a mild form of unruly behaviour is tolerated and visible, social actors tend to modify their behaviour and become disorderly as well. The authors of the broken window theory, James Wilson and James Kelling explain it as follows: “Consider a building with a few broken windows. If the windows are not repaired, the tendency is for vandals to break a few more windows. Eventually, they may even break into the building, and if it's unoccupied, perhaps become squatters or light fires inside” (Wilson and Kelling, 1982). In the context of

¹⁶ Entrenched uncertainty, however, may well make regulatory action difficult. While automation was initially hailed as a solution to market surveillance, the fragmented landscape of global finance may imply that all-encompassing oversight remains a regulatory utopia. As one market participant observed, fragmentation – including the existence of proprietary barriers to information – makes monitoring a thorny task: “if somebody wanted to monitor trading activity across all venues, they would need to combine “attributed” (identifying the owner) and “privileged” (including non-public information on hidden and iceberg orders) data from all the venues. This just isn't possible today.” Interview with Fixnetix executives Hugh Hughes, Chairman and Chief Executive Officer and Bob Fuller, Director, The High-frequency Trading Review.

Impersonal efficiency and the dangers of a fully automated securities exchange

automated trading, the possible dangers resulting from a disorderly market include not only higher volatility or the possibility of other crash-like events, but also network saturation and problems in the time stamping of orders: algorithmic trading strategies rely on a number of measures of network latency that allow traders to estimate the probability of the success of an order entered into the system. Consequently, affecting network latency through techniques such as so-called quote-stuffing may provide advantages to aggressive traders, by providing knowledge of the 'true' state of network latency. The use of such strategies may well lead to an algorithmic arms-race in which actors compete to affect network behaviour in order to artificially control levels of network saturation with respect to the specific execution policies of trading venues.

Pre-emptive behaviour

A variation of the above is a less blatant descent into disorder. After a few years of weak norm enforcement, it will become clear to a majority of market participants that other actors can act opportunistically or contribute to undermining market order. Whether intentionally, because of panic or because of technical incompetence, a rival in weak-norm environment can create a situation that threatens a fund's position. Anticipating this risk, forward-looking actors (even non-opportunistic ones) may be led to start ignoring norms of market order. For instance, anticipating the possibility of broken traders or episodes in market free fall, funds may decide to pull out of the market even more quickly and under a wider set of circumstances, thus contributing to an overall reduction in liquidity.

Overregulation

A related problem in a weak-norms market is the danger of excess state intervention to compensate for the lack of a genuine self-regulatory mechanism. If markets are perceived to be lacking a norm enforcement mechanism, dissatisfaction among the public and politicians could easily lead to regulatory concern and overreaction. We can see an example of this in the current debate about what to do to prevent another Flash Crash. The fourteen recommendations provided by the Joint CFTC-SEC Advisory Committee on Emerging Regulatory Issues aim at providing a mix of standardization and fairness. In a recent post in a specialized electronic forum on market microstructure, a journalist compared the Committee's recommendations with Formula One racing. A similar concern with safety and fairness in Formula One has led to a rigid competition that is dominated by few players. As the post argues, "there are just 12 teams racing in the 2011 season – a 30-year low – and all the cars are powered by just four engine manufacturers. Why? Sixty years of increasing regulation and an accelerating technical arms race have taken their toll." Overregulation in finance can similarly have a negative impact on the market by increasing costs, reducing the number of actors in the market – and thereby the diversity of strategies –and narrowing competition to a few rigid dimensions.¹⁷

In particular, the rigidities induced by overregulation seem to be two-fold. On the one hand,

¹⁷ Another example of overregulation is offered by the American Dodd-Frank act, introduced in response to the credit crisis. This regulation introduces the equities and futures markets model -- along with fragmentation, price transparency, etc. -- into other asset classes, opening the possibility of a flash crash.

Impersonal efficiency and the dangers of a fully automated securities exchange

some trends in the regulation of high-frequency domains seem to indicate an expansion of the bureaucratic control over the financial services industry.

On the other hand, such expansion of bureaucratic control would have an unavoidable cost to end-users (in addition to possible costs associated to information leakage). In a market driven primarily by costs, the imposition of additional regulatory burdens would quite possibly lead to consolidation of intermediaries, leading to the situation outlined above: a financial landscape populated by a handful of 'engine manufacturers'.

Risks stemming from toxic transparency

From a historical perspective, the prices produced by exchanges and other trading venues today are considerably timelier than a generation or two ago. Advances in technology and market organization have made price dissemination almost instantaneous for all non-critical purposes. The speeds at which prices are produced and generated, however, have required a rapid expansion of the mediating technologies that allow for market data to be processed and visualized. This entails three potential risks, described below.

Increased cognitive mediation

The strongly mediated character of price information may be a source of additional future risks. Market prices are not the outcome of a straightforward interpersonal mechanism of market negotiation but, rather, are generated by a variety of impersonal instruments and techniques. This opens the possibility for challenges to the legitimacy of their use and meaning. This change in what could be termed "cognitive mediation" is illustrated by the following historical comparison. When prices were made on the floor of the London Stock Exchange prior to 1986, they existed as utterances -- as quotes generated by a jobber to inquiring brokers. Investors wishing to obtain the best price in the market needed only to phone their brokers, who would collect information from the floor and report back, via the clerks located off the floor, the prices of particular securities. Price information was mediated by people and devices in a way that kept interpersonal connections relatively unprocessed. Today, in contrast, price information is the product of technological intermediaries (Beunza and Stark 2004), including elements of synthetic representation. Since speed makes electronic order books impossible to process with human capacities, technical solutions are necessary in order to present price information to users. No one has, for instance, direct visualization of the order book as-is, requiring the use of some form of technological visualization systems (e.g. Fidessa and Panopticon; see Pryke 2010). One potential danger arising from this necessity is a tighter coupling of the trading process: the more durable are the links between the different steps in the trading process, the more vulnerable is the system to a normal accident.

Speed and the 'reality gap'

The fact that prices are timely can create additional problems of credibility. For latency-sensitive traders, the value of prices is closely associated to their age, often measured in milliseconds. The "younger" a price, the closer it is to the record held on the matching engine and thus the more confidence there is in trading at that specific price. The network problems reported by the NYSE during the Flash Crash, and the problems posed by so-called quote-stuffing highlighted by software specialist firm Nanex, are two examples of how the representativeness of prices may be hampered by technological conditions.

Importantly, the 'reality gap' between the order book and the user systems, as latency is lowered throughout the system, will probably lead to new layers of technical intermediation by

means of which market participants make prices meaningful. An example of how such gap leads to technological fragmentation is provided by the systems of an important London-based market maker¹⁸. The firm, which is engaged in a variety of markets, is faced with a varied set of technical specifications and order-matching rules. In this context, delayed confirmations of either quote modifications, order cancels or trade executions from any single venue may result in unwanted fluctuations in the firm's inventory levels. Such delays, however, are contingent on the venue's quote stamping policies, order-matching rules and network topology. The chaotic nature of network latency, coupled with the fast pace transactions turn the availability of counterparties into uncertainty.

To increase the chances that orders are met with willing counterparties, the firm simulates the rules and conditions on all the trading venues they are engaged.¹⁹ On submitting a cancel order and not receiving a confirmation they may, for instance, calculate the probability that the said order was received by an exchange, given the state of the network and the trading venue's rules. Upon this calculation, decision algorithms may send new orders to keep the firm's planned market position. This simulation acts as an intermediary between the decision algorithms (which decide on strategies based on the input parameters of traders) and the matching engines of the market. The firm's algorithms, in other words, do not 'see' the actual market, but a simulation of it.

One potential problem with so-called epistemic accidents is the difficulty in handling them on the part of the regulator. By their nature, epistemic accidents do not lend themselves to consensus about "what happened." This, in turn, might make it difficult for regulators to justify drastic action – something akin to the British or American intervention in financial institutions during the credit crisis. Thus, a highly automated system may not only be more prone to accidents, but more polemic in its solutions.

Challenges of non-price information

Prices and trading volume have always been affected by a myriad of informational elements, and low-latency markets are not different. However, the crucial distinction between the status of information in markets before the advent of high-frequency trading is that the latter require a much broader array of knowledge.

To explain this challenge we need to understand the material nature of profit-formation in the high-frequency trading world. The underlying principle of profit, as the name implies, is speed. Speed calls for new dimensions of performance that require their own specialized metrics. Developing theories to grasp high-frequency markets calls for expanding the analytical scope of economics into computer science and networks engineering. The factors that affect the speed at which an order can be placed are, among other things, the topology of the possible routes between the order-originating computer and the exchange's order-matching computer; the overall physical length of cables connecting the two computers; the level of traffic on the network and the temperature of the network, as this impacts the electrical resistance of the different components in the network. High-frequency traders take all of these factors (and many others not listed here) into account when setting their trading algorithms. However, there exists no economic theory that incorporates these factors into its analytical worldview.

¹⁸ Interview with HFT software developer, London, February 2011.

¹⁹ It is interesting to note that some of the techniques used in the design of this system resemble those deployed to make latency-sensitive decisions on massively multiplayer online role-playing games.

Potential risks stemming from fragmented innovation

Technical failure and the erosion of institutional confidence

Under fragmented innovation, knowledge-intensive competition will develop, which will make profits contingent on the production and control of proprietary knowledge. The algorithms associated with high-frequency statistical arbitrage are an extreme example of this, having a half-life of between one day and two weeks, as they are often discovered by competitors through reverse-engineering strategies. These short time frames reveal the sensitivity of information that characterizes today's markets. It is therefore not surprising that fragmentation has resulted in an overall disincentive for cooperation, leading participants to develop a 'tunnel vision' of the financial landscape.

The limits to innovation in this environment are not clear. For instance, some prominent members of the industry believe that intense competition will lead to a situation in which high-frequency trading will cease to be profitable²⁰. Indeed, it is quite possible that economies of scale will drive high-frequency activity to a handful of large market participants engaged primarily in electronic market making and forms of latency-sensitive statistical arbitrage. However, preliminary evidence suggests that the growing number of lit and dark trading venues, along with the large number of market participants and the increased interconnections across markets make of algorithmic trading not a momentary trend but rather a defining feature of future markets. It is important to note that the constant co-evolution of physical infrastructures and financial instruments in the market generates a continuously expanding horizon of possibilities for deploying algorithms in the market. Thus, there is no visible limit to the number of algorithms that can be deployed in the market and, consequently, to the speed of innovation in this field.

One consequence is the increased complexity faced by established exchanges. First, the innovation pipelines of large exchanges such as the LSE are not necessarily aligned with those of their leading users. Regulated exchanges have longer technological investment horizons than smaller organizations²¹. Second, established venues face a higher level of interconnection complexity. The decentralization and globalization of trading means that these institutions must now negotiate the operational parameters of new systems with their customers. Importantly, however, their systems remain the key nodes of the financial infrastructure. A parallel exists, in this sense, between the development of operating systems and the problems faced by trading venues. The failure of Microsoft Vista illustrates the difficulties of coordinating the technological evolution of several technologies, when all of them need to fit together. Microsoft needed to design an operating system that would allow for plug-and-play functionalities for a variety of computer peripheral vendors. At the same time, the company also needed to rely on estimates for the evolution of microprocessor speed. But peripheral manufacturers did not improve their performance by as much as microprocessors did. The operating system was caught up in the gap, suffered from poor performance and ended up a commercial failure.

²⁰ Interview with former CEO of a Regulated Market, April 2011.

²¹ Interview with Technology Manager of a regulated market, London, February 2011.

Impersonal efficiency and the dangers of a fully automated securities exchange

Taken together, these factors may well lead to an increase in the number of outages in key trading venues – in the same way that added complexity, and compressed design timelines may have contributed to the instability of the maligned Microsoft Vista. Trading outages are the outcome of several factors. First, increased interconnection complexities between trading venues and the systems of market participants; second, a pressure to roll out new services to meet perceived user demands; and third, an increased opacity in terms of how a new system will change the technical landscape of the marketplace. While so-called “dress rehearsals” may provide some information on systems before they are widely introduced, they are simply unable to replicate all of the conditions that lead to outage and/or systemic failure.

To date, the strategy in relation to outages has been to either develop more robust IT management procedures or follow transparent guidelines on how to act in the event of failure. The FSA, for instance, presents good practice as maintaining open channels of communication between the venue, its participants, and the market, and keeping restart as non-discretionary as possible (see FSA 2010). Such view considers outages as relatively unavoidable, though manageable events, which do not threaten the role of key trading venues as centres of price formation.

The role of established trading venues as centres of price formation, however, hinges on a small set of public routines that are deemed authoritative and widely reliable by market participants. The opening auction at the London Stock Exchange is such an example, setting the opening and closing prices for the stocks in the market. Outages among the less adaptable organizations thus carry the risk of a continued erosion of confidence that may result in a less structured, more disparate price formation system. The continued pressure towards decentralization induced by the erosion of the current price formation centres may well result in a diminished capacity to raise capital in the primary market. Unlike regulated markets, multilateral trading facilities – and indeed, much of the activity in electronic markets – are oriented towards the provision of services for the secondary market.

Risk relief mechanisms

In addition to risks, automation can also bring about opportunities that reduce the risks identified above. One possible direction is computer-enhanced sociability, currently visible outside the realm of finance in the shift within new media towards the Web 2.0. Just as Facebook differs from a regular website in that it promotes social engagement and community-building, new and more advanced trading systems may well provide markets with functionalities for social interaction, social cues and human judgment²². We see this happening in two different ways.

²² Contemporary financial markets are populated by numerous instances of computer-enhanced sociability, a fact that the authors of this report recognize. Nonetheless, it is important to recognize that such sociability is often limited to certain communities (e.g. traders) who may not have the position necessary for reflecting on the risks accrued within the system. For instance, traders often make use of instant messaging systems and specialized forums to communicate with their peers; but such systems are ultimately aimed at exchanging information that is not overtly associated to the type of sociotechnical risks discussed in this paper. Thus, while a group of traders may communicate via electronic systems on the apparent value of a certain firm or on the contents of a specific research document, they are unlikely to engage in a reflexive analysis of the sociotechnical risks of the market at a particular point in time. Indeed, such communication systems may lead users to mimic each other's behaviours, leading to the creation of risky consensus positions in the market.

Bringing back the social dimension

At the level of exchanges, the incorporation of “the social” into financial algorithms may bring back the benefits of intermediaries. This process has been studied by French sociologist Fabian Muniesa (2007). In his analysis of the automation of the closing auction of Paris Bourse, Muniesa, uses a culinary metaphor to describe how automation can retain a social component: according to him, just as cocoa powder has to be carefully “folded” into fresh cream to preserve the qualities of the cream, the social aspects of a live trading floor need to be understood and then adapted into algorithms to preserve the social component of a market. Muniesa shows how different algorithmic rules (average of last trades, last trade) satisfied to a different degree competing goals of a closing auction (fairness, stability, robustness to manipulation). The implication is that that the move towards electronic trading entails explicit choices about what type of market is desirable. In other words, automation does not necessarily imply impersonality.

Nevertheless, the practice of so-called “folding” requires exchanges to overcome substantial barriers between technically sophisticated programmers and people-oriented brokers and market makers. Following the Flash Crash, the recognition that an asocial market is problematic may prompt renewed energies in this direction. Indeed, in his analysis of the Flash Crash, the economic historian Paul David (2010:3) outlines the elements that such market should have: it needs, according to him, “some substantial minimum degree of mutual recognition of the participants’ common interest, enough to encourage reflexive identification of individual actors with the group.” Only such a system, David argues, can induce compliance with norms and the necessary social control to stabilize interactions.

Bringing back human judgment

At the level of traders, individual human judgment could stage a comeback with the creation of interfaces that allow actors to exercise discretionary judgment, even in an automated environment. A market maker’s decision to pull out of the market is a matter of judgment. With the rise of faster markets, much of this decision has been incorporated into the algorithms in the form of automated hard-and-fast rules. But there are alternatives if better interfaces can be designed. In Beunza’s interviews during 2009 with designers of financial interfaces – handheld computers, market maker’s screens – he was struck by the depth of their thoughts in this respect. One designer, for instance, drew an analogy between automating markets and manufacturing. Just as replacing three workers by a machine does not eliminate the possibility for human control of the process – control simply moves to the knobs and levers in the machine – neither does automated trading have to mean relinquishing human control. The digital equivalent of knobs in a trading interface may give ever-greater control to the traders, and afford the possibility to use their judgment in an automated setting. In this respect, it is interesting that the telecom company BT has recently been researching the use of the iPad in hedge funds. The company, which runs a unit specialized in financial services, gave iPads to fund managers and studied its use for two months with surveys. The introduction of control in automated systems will remain an important challenge to policy-makers, who will have to trade off costly human judgment for efficient, yet potentially risky, rule-based automation.

Learning from instances of market stress

Aside from a re-socialization of markets, the occurrence of more instances of market stress may reduce in the probability of a large crisis. Automated traders, and especially those in high-frequency trading, operate in considerable secrecy. The result is fragmented innovation. But diagnosing and fixing situations of market stress can address this fragmentation. For instance,

Impersonal efficiency and the dangers of a fully automated securities exchange

one observation from Beunza's interviews with high-frequency traders after the Flash Crash is that the crisis gave traders a better sense of where their competitors stood. In a marketing event by a large European exchange in London, the Exchange discussed the capabilities offered by his organization, with special reference to the Flash Crash. Based on the presentation, a high-frequency trader that also attended the event found out that his own practice of switching off one data communication channel to increase speed was something that other competitors did as well. In short, because crises prompt sense making and new offerings from exchanges, every incident involves the possibility of learning from them -- but one cannot guarantee that this will happen.

References

- Baker, W. (1984). 'The Social Structure of a National Securities Market,' *American Journal of Sociology*, 775-811.
- Barnes, R. (2008). 'The race is on,' *Focus: The Market Magazine*. October: 44-53
- Beunza, D. and D. Stark (2004). 'Tools of the trade: the socio-technology of arbitrage in a Wall Street trading room.' *Industrial and Corporate Change* 13(2): 369.
- Black, Fischer (1971). 'Toward a fully automated stock exchange, part I.' *Financial Analysts Journal*, 27 (4), p. 28.
- Blume, M. E., J. Siegel and J. Rottenberg (1993). *Revolution on Wall Street: the rise and decline of the New York Stock Exchange*. Norton & Company: New York and London.
- Christie, W. G. and P. H. Schultz (1994). 'Why do NASDAQ Market Makers Avoid Odd-Eighth Quotes?' *The Journal of Finance*, 49(5), p.1813.
- 'Economic Links and Predictable Returns.' *The Journal of Finance*, 63:1977-2011.
- Cohen, L., A. Frazzini, and C. Malloy. (2010). 'Sell-Side School Ties.' *The Journal of Finance*, 65:1409-1437.
- David, P. A. (2010). 'May 6th – Signals from a Very Brief but Emblematic Catastrophe on Wall Street.' Working Paper. <http://ssrn.com/abstract=1641419>
- Domowitz, I. (1993). "A taxonomy of automated trade execution systems." *Journal of International Money and Finance*, 12 (6): 607.
- Downer, J. (2010a) 'Anatomy of disaster: why some accidents are unavoidable' *CARR Discussion Paper* 61. Available at <http://www2.lse.ac.uk/researchAndExpertise/units/CARR/pdf/DPs/anatomyOfDisaster.pdf>
- Ellis, N., L. Fairchild and H. Fletcher (2009) 'The NYSE response to specialist misconduct: an example of the failure of self-regulation,' Unpublished paper (http://works.bepress.com/cgi/viewcontent.cgi?article=1000&context=nan_ellis, accessed on 8.3.2011).
- Gasparino, C. (2007) *King of the Club: Richard Grasso and the Survival of the New York Stock Exchange*, New York: Collins Business.
- Hong, H., J. D. Kubik, and J. C. Stein. (2005). "Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers." *The Journal of Finance* 60:2801-2824.
- Financial Services Authority (2010). Market Watch, Issue 36, August.
- Greenwald, B. and J. Stein (1988). 'The task force report: the reasoning behind the recommendations.' *The Journal of Economic Perspectives*, 2(3), p. 3.
- Lessig, L. (1999). *Code and Other Laws of Cyberspace*. Basic Books. New York.

Impersonal efficiency and the dangers of a fully automated securities exchange

- Levinthal, D. A., and J. G. March. (1993). 'The myopia of learning.' *Strategic Management Journal* 14:95-112.
- Lucchetti, A. (2007). 'Big Board's next chief, ex-trader at Goldman, may make more deals.' *The Wall Street Journal*, 15 November 2007, p. A1.
- MacKenzie, D. (2006). *An Engine, Not a Camera. How Financial Models Shape Markets*, Boston: MIT Press.
- MacKenzie, D. and Millo, Y. (2003). 'Constructing a Market, Performing Theory: The Historical Sociology of a Financial Derivatives Exchange', *American Journal of Sociology*, 107-145.
- March, J. G. and Z. Shapira (1987). 'Managerial Perspectives on Risk and Risk Taking' *Management Science*, 33:1404-1418.
- McNamara, G. and Philip Bromiley (1997). 'Decision Making in an Organizational Setting: Cognitive and Organizational Influences on Risk Assessment in Commercial Lending' *The Academy of Management Journal* 40: 1063-1088.
- Michie, R. C. (2001). *The London Stock Exchange: A History*. Oxford University Press: Oxford.
- Millo, Y. and D. MacKenzie (2009). 'The usefulness of inaccurate models: Towards an understanding of the emergence of financial risk management.' *Accounting, Organizations and Society* 34: 638-653.
- Mirowski, P. (2002). *Machine dreams economics becomes a cyborg science*, Cambridge University Press: Cambridge.
- Mittal, H. (2008). 'Are you playing in a toxic dark pool? A guide to preventing information leakage,' *Investment Technology Group*, June 2008. Available at http://www.positjapan.org/news_events/papers/ITGResearch_Toxic_Dark_Pool_070208.pdf., Accessed on 8.3.2011.
- Muniesa, F. (2007). 'Market technologies and the pragmatics of prices.' *Economy and Society*, 36(3): 377-395.
- Nanex, 2011. 'The Mini Flash Crash of 04-28-2010' Available at http://www.nanex.net/FlashCrash/FlashCrashAnalysis_042810_MiniFlashCrash.html Accessed on December 10, 2010.
- Pardo-Guerra, J. P. (2010). 'Creating flows of interpersonal bits: the automation of the London Stock Exchange, 1955-1990.' *Economy and Society*, 39(1), p. 84.
- (2011). 'The automated house: the digitalization of the London Stock Exchange, 1955-1990' in *Technological innovation in retail finance: international historical perspectives*. B. Batiz-Lazo, J. C. Maixé-Altés and P. Thomes. Routledge: London.
- Perrow, C. (1999). *Normal accidents: living with high risk technologies*. Princeton University Press: Princeton, NJ.
- Pryke, M. (2010). "Money's eyes: the visual preparation of financial markets." *Economy and Society*, 39(4), p. 427.

Impersonal efficiency and the dangers of a fully automated securities exchange

- Securities and Exchange Commission (1963). *Report of the Special Study of the Securities Markets of the Securities and Exchange Commission* (88th Cong., 1st sess., House Document 95, 1963) Government Printing Office: Washington, D.C.
- (1971). *Institutional Investor Study Report of the Securities and Exchange Commission*, (H.R. Doc. No. 64, 92d Cong., 1st Sess.) Government Printing Office: Washington, D.C.
- (1974). 'Notice of SEC Action Declaring Effective a Consolidated Tape Plan Filed Pursuant to Rule 17a-5 and the Securities Exchange Act of 1934,' (May 10th), 39 Fed. Reg. 17799.
- (1978a). *Dissemination of Quotations for Reported Securities*, (January 26th), 43 Fed. Reg. 4342.
- (1978b). *Development of a National Market System*, 43 Fed. Reg. 4354 (February 1st).
- (1978c). 'Securities Exchange Act Release No. 34-19456,' 48 Fed. Reg. 4938 (February 3rd).
- Seligman, J., (2003). *The transformation of Wall Street*, Aspen Publishers: Gaithersburg.
- (1985). *The SEC and the future of finance*, Praeger Publishers: Westport, CT.
- Shapiro, S. P. (1987). *Wayward capitalists: targets of the Securities and Exchange Commission*, Yale University Press: New Haven.
- Simon, M. and Susan M. H. (2003). 'The Relationship between Overconfidence and the Introduction of Risky Products: Evidence from a Field Study'. *Academy of Management Journal*, 46: 139-149.
- Stein, Jeremy C. (2008). 'Conversations among competitors.' *American Economic Review* 98(5): 2150-2162.
- Thomas, W. A. (1973). *The provincial stock exchanges*. Cass: London.
- US Treasury (1988). 'Report of the Presidential Task Force on Market Mechanisms: submitted to The President of the United States, The Secretary of the Treasury, and The Chairman of the Federal Reserve Board.' The Task Force: Washington, D.C.
- Wilson, J. and Kelling, G. (1982). 'Broken Windows,' *Atlantic Monthly*, March (249): 29-38.
- Wunsch, S. (2011). 'Straitjacket' (parts 1-3) Available on <http://www.tabbforum.com/opinions/> Accessed on 9.1.2011.
- Zaloom, C. 2003. 'Ambiguous Numbers: Trading Technologies and Interpretation in Financial Markets'. *American Ethnologist*, 30(2): 258-72.
- Zuckerman, E. (2004). 'Structural Incoherence and Stock Market Activity.' *American Sociological Review*, 69: 405-432.

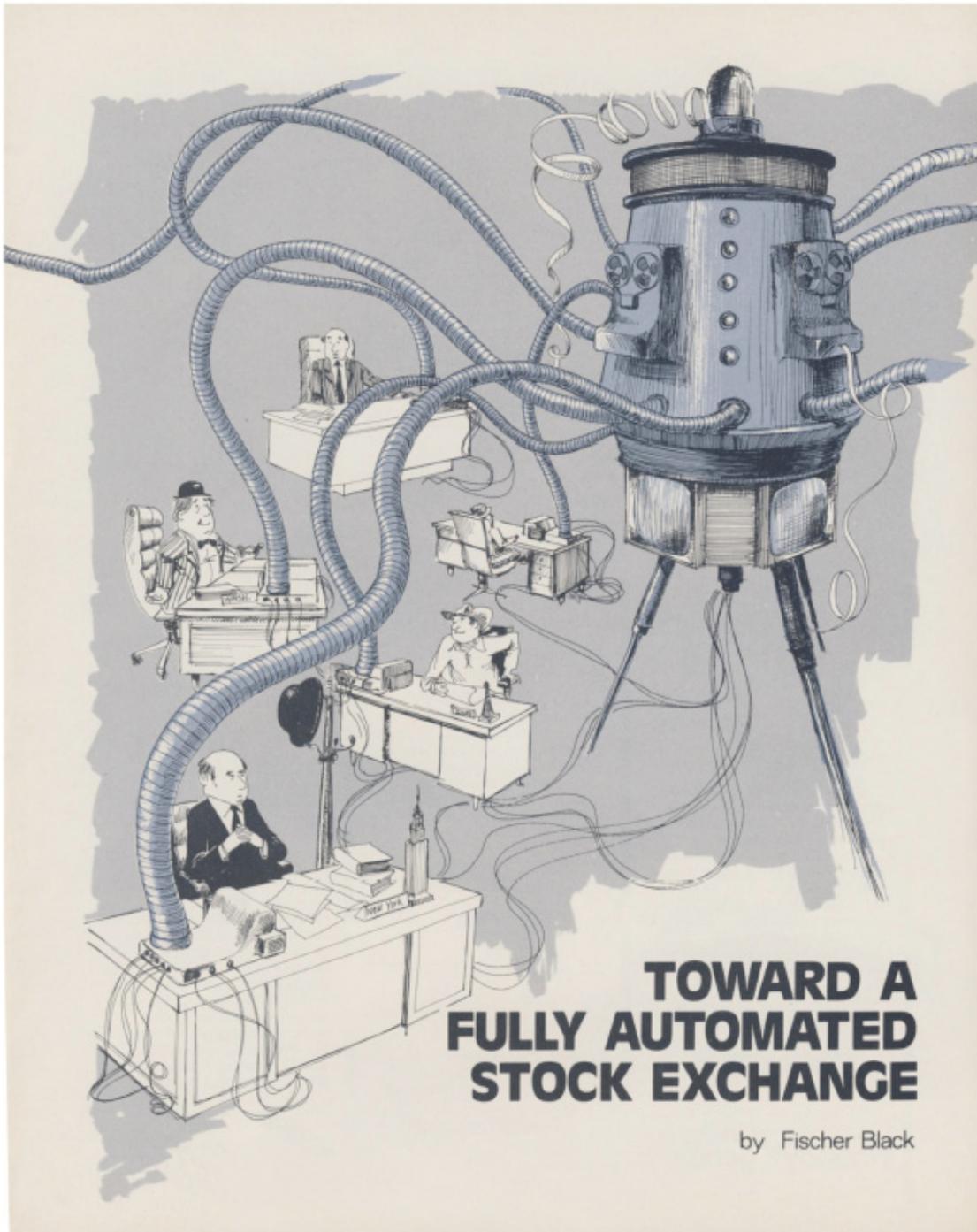


Figure 1. Illustration of Fischer Black's article (1971).

Impersonal efficiency and the dangers of a fully automated securities exchange



Figure 2. The markets of the future, as seen by the Strategic Engineering Unit of the London Stock Exchange (c. 1986)

