

**Drilling Through the Allegheny Mountains:  
Liquidity, Materiality and High-Frequency  
Trading**

**Donald MacKenzie**

**Daniel Beunza**

**Yuval Millo**

**Juan Pablo Pardo-Guerra**

January 2012

**Abstract**

In 1999, Carruthers and Stinchcombe provided the classic discussion of ‘the social structure of liquidity’: the institutional arrangements that support markets in which ‘exchange occurs easily and frequently’ (1999, p. 353). Our argument in this paper is that the material aspects of these arrangements – and particularly the materiality of prices – need far closer attention than they normally receive. We develop this argument by highlighting two features of new assemblages that have been created in financial markets since 1999. First, these assemblages give sharp economic significance to spatial location and to physical phenomena such as the speed of light (the physics of these assemblages is Einsteinian, not Newtonian, so to speak). Second, they have provoked fierce controversy focusing on ultra-fast ‘high-frequency trading’, controversy in which issues of materiality are interwoven intimately with questions of legitimacy, particularly of fairness.

**KEYWORDS:** social studies of finance; high-frequency trading; liquidity; materiality.

By liquidity of a market, economists mean that standardized products can be bought and sold continuously at a price that everyone in the market can know ... The idea is that everyone can know at all times what the price is, and only one price obtains in the market. (Carruthers and Stinchcombe 1999, p.353)

[W]e have to abandon this idea that there is a universal truth for the best currently available price. (Natan Tiefenbrum, Commercial Director of the electronic trading venue Turquoise, interviewed by *High Frequency Trading Review*, 10 Dec 2010)

'Liquid markets' are a powerful metaphor and a rare social achievement.<sup>1</sup> Carruthers and Stinchcombe (1999) suggest that 'three basic mechanisms' make liquidity possible: first, a continuous competitive auction; second, market makers 'who for a small margin, are willing to take the risk of transferring large quantities and maintain a continuous price'; and third, the creation of homogeneous and standardized commodities, such as 'legal instruments with equal claims on an income stream' (1999, p. 353).

As Carruthers and Stinchcombe point out, for buyers and sellers to be ready to trade with each other at a price they can quickly agree on, they need 'to know the commodities they transact in', and liquidity is thus 'an issue in the sociology of knowledge' (1999, p. 353). Implicitly, therefore, their article builds a bridge between economic sociology and the new 'social studies of finance' (only nascent in 1999), which, inter alia, brings to bear perspectives from science and technology studies (STS). The core viewpoint of STS-inspired research on finance (a viewpoint developed in Callon 1998 and in much subsequent work by him and others) is that economic actors are sociotechnical ensembles, not 'naked' human beings. (Amongst the foundations of this viewpoint is actor-network theory, so we were delighted to discover that one Swiss firm that supplies software for automated trading is

---

<sup>1</sup> For the metaphor, see the August 2011 issue of this journal; for the argument that even stock markets 'are normally illiquid with temporal spikes of liquidity', see Pitluck (2011, p. 26).

actually called Actant: see [www.actant.com](http://www.actant.com).) Actors' 'equipment' matters. Amongst the themes in this literature – again, a theme initially inspired above all by Callon (1998) – is that economic models are not simply representations, but interventions in markets: part of the processes by which markets are constructed. However, this claim (that economics is 'performative') is simply one aspect of a more general focus on the *material* nature of markets: their physical, technological, corporeal nature.

Carruthers and Stinchcombe concentrate on the third of their 'basic mechanisms', on what they call the 'minting' work needed to create homogeneous commodities. Our paper focuses on the first two, continuous auctions and market making. We examine the radical shift since 1999 in their technological underpinnings, which has made possible and also been reinforced by the rise of automated trading, particularly fully automatic, ultra-fast 'high-frequency trading'. The changing material assemblages that constitute 'liquid' markets deserve detailed attention, we argue - especially the materiality of prices (for which see, e.g., Muniesa 2007). A price is not an abstraction: to be conveyed from one human being to another, or from one automated system to another, a price must take a material form, whether that be the sound waves created by speech, the electrical signals of the telegraph or telephone, or the optical signals that now flow through high-speed networks. As Carruthers and Stinchcombe note, 'the agreement of a large number of buyers and sellers that a given commodity has a given price, at which they could all buy it or sell it, is central to liquidity' (1999, p. 379). As we shall discuss, the changing material assemblages of liquid markets have rendered that 'agreement' precarious, as the second quotation with which we began suggests.

While the literature in social studies of finance contains excellent studies of the automation of stock exchanges (including Muniesa 2005 and 2007 and Pardo-Guerra 2010) and of manual ('point and click') trading via automated exchanges (notably Zaloom 2006 and Preda 2009), it is only just beginning to encompass automated trading. We know of only one paper focused directly on this – Lenglet (2011), a fascinating observational study of

the ‘execution algorithms’ discussed below and of the regulatory issues they raise – although Muniesa (2007), Beunza and Stark (2004) and Beunza et al. (2011) also discuss automated trading in the context of, respectively, price formation, arbitrage, and the ‘ideology of impersonal efficiency’. The economics literature is larger: see e.g. Brogaard (2010), Chaboud et al. (2009), Hasbrouck and Saar (2010), Hendershott et al. (2011), and Jovanovic and Menkveld (2010). Even it, though, has limitations, mainly because quantitative data are limited: even though some (by no means all) trading venues require orders to carry a digital identifier indicating whether an order was generated by a human or by a machine, such identifiers are not normally public. Therefore, for example, the TABB Group market-share data drawn on below are estimates based on triangulating published data on overall trading volumes with interviews with market participants.

This article is based on three sets of sources. First is 19 exploratory interviews with traders, brokers, analysts, managers of electronic exchanges, etc., mainly in London and New York. Second is observations and informal conversations with participants in High Frequency Trading World Europe 2011, the field’s main European meeting, and a visit to the command centre of Globex, the electronic trading system of the Chicago Mercantile Exchange. Third is documentary sources, including the relevant trade press (such as the magazine *Automated Trader*), contributions to the debate on high-frequency trading (e.g. on the TABB Forum, the liveliest meeting point between proponents and opponents of high-frequency trading), and such limited literature as exists on the techniques of automated trading (e.g. Durbin 2010). Although our research also encompasses developments in Europe, for reasons of space we focus in this article primarily on the U.S.

Four sections follow this introduction. First, we discuss electronic public-limit order books, which are the chief material form that Carruthers and Stinchcombe’s continuous auctions now take, and automated trading itself. We outline the latter’s main types, including the most important from the viewpoint of the new ‘social structure of liquidity’: electronic market-making. The second section discusses issues of time, space and the materiality of

prices: while it is tempting to believe that globalized financial markets have brought about ‘the end of geography’ (O’Brien 1992), high-frequency trading gives the obdurate physical reality of space a renewed prominence, and a physical constraint – the speed of light – is of growing importance. The third section discusses the fierce controversy that has erupted since late 2008 over the legitimacy of high-frequency trading; the section focuses on the contested question of the ‘fairness’ of high-frequency trading, and we examine specific ways in which ‘fairness’ and materiality are entangled. The fourth section, the article’s conclusion, asks just how novel is the new ‘social structure of liquidity’ to which automated trading gives rise.

## **1. Electronic Public Limit-Order Books and the Rise of Automated Trading**

As Carruthers and Stinchcombe (1999) suggest, liquidity in financial markets traditionally involved a distinct (often officially designated) category of human participant, market makers. They received buy or sell orders from other market participants, manually matched them and/or acted as dealers, continuously quoting prices at which they would themselves buy from or sell to other market participants. From the late 1980s onwards, however, ‘human-mediated’ markets of this kind have increasingly been challenged by electronic public limit-order books. (A ‘limit order’ is an order to buy at or below a given price, or alternatively an order to sell at or above a given price.) Market participants can place orders in the book directly, and the book is public in the sense that it is visible to all market participants (but not normally to the general public). A ‘snapshot’ of a simple, hypothetical book is shown in figure 1. (Actual public limit-order books are fast-changing and sometimes complex, and sophisticated visual-representation software is sometimes used to help human traders grasp them: see Pryke 2010.) In figure 1, the book contains offers to sell 200 shares at \$21.00, 150 shares at \$21.01, etc., and bids to buy 100 shares at \$20.99, 450 shares at \$20.98, etc.

— FIGURE 1 AROUND HERE —

Public limit-order books are maintained on computer systems known as 'matching engines' because they execute a trade when they find a bid to buy and an offer to sell that match. (In the book shown in figure 1, there is no match. However, a match could be created by a market participant entering a bid to buy at \$21.00 or an offer to sell at \$20.99, or by the arrival of a 'market order': an order simply to buy or sell at the best available price.) Public limit-order books were generally created first at the margins of the global financial system, such as by the Toronto Stock Exchange or the new Swedish electronic options exchange, OM (Optionsmäklarna), set up in 1985 (Muniesa 2005, Gorham and Singh 2009). Market makers on the world's main exchanges (such as the 'specialists' of the New York Stock Exchange or Chicago's open-outcry pit traders) saw electronic trading as a threat to their profitable business. In those exchanges, such trading was for many years generally restricted to small retail orders and/or times when the trading pits were closed.

During the early 2000s, however, resistance by traditional market makers crumbled, as established exchanges faced increasing competition from fully-electronic counterparts, such as Eurex (which traded futures entirely electronically: see Scott and Barrett 2005) and Island, a trading venue<sup>2</sup> established in New York in 1997 to allow market participants directly to trade Nasdaq-listed stocks without having to do so via Nasdaq's market makers. Aspects of all three of the 'mechanisms of institutional isomorphic change' on which DiMaggio and Powell (1983, p.50) focus – coercive, normative and mimetic – were in play. In the U.S., decisions in the 1990s and 2000s by the chief stock-market regulator, the Securities and Exchange Commission (SEC), helped foster electronic trading. The Securities Exchange Act of 1934 laid down the goal of 'to remove impediments to and perfect the mechanism of a free and open market and a national market system', and that perfected 'mechanism' was increasingly seen, by the SEC and others, as an electronic

---

<sup>2</sup> Such venues include ECNs (electronic communication networks), such as Island, which are officially recognised but are not stock exchanges (i.e. do not have the power themselves to 'list' companies).

public limit-order book.<sup>3</sup> Electronic trading in which external participants could place their orders directly was frequently perceived as fairer to them than older forms of trading in which they had to place orders via human intermediaries who (as we touch on in the conclusion) had economic interests that were not always aligned with those of their customers. By the 2000s, indeed, an electronic public limit-order book was often seen simply as a marker of a 'modern' securities exchange.

An electronic public limit-order book does not necessitate automated trading. Orders originally exclusively were (and many still are) placed 'manually' by human traders using mouse and/or computer keyboard (see Zaloom 2006 and Preda 2009). However, the situation Zaloom describes shows how an electronic public limit-order book provides an incentive for automated trading: the opportunities she, as a participant observer, and her fellow traders were exploiting were fleeting, and for human traders to capture them required fast reactions and excellent hand-eye coordination. Automation, furthermore, was clearly technically feasible, in that placing an order had become simply sending an electronic message from a computer terminal. There were technical problems to overcome – in particular, software interfaces between exchanges' matching engines and trading algorithms needed to be developed – but they were far from insuperable.

The resultant world of automated trading contains two main camps. One camp is 'proprietary-trading' systems that are the automated equivalents of Zaloom's colleagues' efforts to make money by trading: we turn to these below. The other camp is 'execution algorithms' used by institutional investors such as mutual funds, pension funds and insurance companies (for which see Lenglet 2011). As Pitluck (2011) shows, these big institutions cannot easily find 'liquidity'. Large orders usually have considerable adverse 'market impact': prices increase as a big investor buys, and fall as it sells. Institutional traders (or brokers acting on their behalf) have therefore long sought to break up large orders into smaller parts and to execute those parts

---

<sup>3</sup> Securities Exchange Act of 1934, section 15A, paragraph b6 (<http://taft.law.uc.edu/CCL/34ACT>, accessed 26 August 2011).



at favourable times (Pitluck 2011 describes this as ‘distributed execution’). A typical large U.S. institutional order might involve 200,000 shares, while the average size of trades executed on Nasdaq or the New York Stock Exchange is now less than 300 shares (<http://liquidnet.com>, accessed 14 September 2011).

Execution algorithms automate the human ‘distributed execution’ described by Pitluck. The first generation of them (which simply split large orders into equal-sized parts, and entered them into order books at equally-spaced times: Leinweber 2009, p. 74) were easy for other traders to spot, and involved no effort to choose favourable times. They have therefore been replaced by a second generation that (though also often predictable) seek to optimize execution times. Such algorithms include ‘volume participation’ programs, which track the number of the securities in question traded in the immediately preceding time period, and submit an order that is a set proportion of that volume (in the hope that market impact will be less when volumes are high), and ‘volume-weighted average price’ or VWAP programs, which seek to anticipate trading volume, for example based on the way volumes typically rise and fall somewhat predictably at different times of the day. A third generation of execution algorithms, now entering into widespread use, are non-scheduled ‘liquidity seekers’, which, in the words of an interviewee, ‘go out and get [liquidity], and hide at the same time’.

Ranged against execution algorithms – and that, interviewees reported, is how it is often seen – is the other camp: proprietary-trading algorithms (it is from these programs that the interviewee quoted in the previous paragraph believes sophisticated execution algorithms need to ‘hide’). Our interviews and documentary research suggest that proprietary algorithms can in principle be grouped into five main categories. First is electronic market-making, the automated equivalent of human market makers’ efforts to earn their ‘small margin’ (Carruthers and Stinchcombe 1999, p. 353) by buying securities at the best bid (in the example in figure 1, \$20.99) and selling at the best offer (in that example, \$21.00). A market-making algorithm continuously quotes a price at which it will buy and a higher price at which it will sell, to try to earn

the bid-offer 'spread' (one cent per share in this example) between the two. As the name suggests, electronic market-making predominantly involves 'providing liquidity' (posting limit orders that others execute against): one market maker estimates that this forms 'approximately 80% of the firm's executed volume' (GETCO Europe 2010, p.1).

A second category is arbitrage between markets, for example exploiting differences between the prices of the same shares quoted in different markets or between equity derivatives (such as stock-index futures) and the underlying shares. A third category is statistical arbitrage, which involves identifying persistent patterns amongst prices (e.g. between pairs of related shares, such as Coca Cola and Pepsi), and betting that prices will return to these patterns if they have diverged from them. A fourth category is order-anticipation strategies (sometimes known as 'algo-sniffing') that seek to identify and exploit execution algorithms whose already-executed trades are simply the tip of a hidden, yet-to-be-executed larger 'iceberg' (see Durbin 2010, pp. 66-68, on how to profit by 'towing an iceberg'). The fifth set of strategies, 'momentum ignition' (SEC 2010, pp. 3609-10) is – if detected – illegal: it involves placing large numbers of orders (the vast majority of which will be cancelled before being executed) with the intention of initiating a rise or fall in prices and profiting from it.

Some of these strategies (above all, electronic market-making) require 'high-frequency trading' or HFT, which involves large volumes of transactions, frequent, rapid placing and cancelling of orders as market conditions change, liquidating positions very quickly (the founder one important HFT firm, Kansas City-based Tradebot, is reported to have said in 2008 'that his firm typically held stocks for 11 seconds': Creswell 2010), and ending the trading day 'flat', i.e. with no inventory of shares or net exposure to price rises or falls. HFT's share of equity trading in the United States was only around 26% as recently as 2006; by 2009, it had risen to about 61 percent: see figure 2 (the

subsequent modest fall is discussed briefly in a note<sup>4</sup>). As table 1 shows, around three-quarters of HFT in the U.S. in 2010 (the last full year for which we have data) was electronic market-making, and because of the counting rule involved (see the caption to figure 2), it is likely that electronic market-making algorithms are involved in most share transactions in the U.S. While there are no directly comparable data on electronic market-making in Europe, HFT's share of European equities trading has been growing fast, and here too electronic market-making is almost certainly by far the leading sector.

— TABLE 1 AND FIGURE 2 AROUND HERE —

Our interviews and document-based research suggest that HFT is conducted by major investment banks (such as Goldman Sachs) and large hedge funds (such as Chicago-based Citadel) but also, very importantly, by specialist firms such as Chicago-based GETCO (Global Electronic Trading Co., the most prominent electronic market-maker), Tradebot and Amsterdam-based Optiver. (To a surprising extent, the origins of HFT in shares seem to lie less in stock exchanges than in derivatives exchanges. Thus GETCO's founders were Chicago pit traders, Tradebot's founder traded on the Kansas City Board of Trade, and Optiver seems to have had its origins in options trading.)

HFTs, and especially electronic market-makers, have become crucial to the provision of liquidity. The most widely used measure of liquidity is the 'spread' between the highest bid to buy and lowest offer to sell (one cent in figure 1). Spreads in the U.S. have come down rapidly in recent years, often falling to the lowest possible value in the current pricing regime, one cent. (The Securities and Exchange Commission regulates the units of price on U.S. public stock trading, and currently one cent is the minimum admissible increment for stocks priced at \$1 or above.) Although multiple factors are involved – most obviously the reduction in 2001 in minimum increment from

---

<sup>4</sup>As the effects of the 2008 banking crisis began to wane in the latter half of 2009, trading volumes and volatility – factors that typically boost HFT profits – went down. Some HFT firms withdrew from the market, an interviewee told us, while others seem to have adjusted their systems to reduce risk, which involves trading less frequently.

one-sixteenth of a dollar to one cent – the growth of electronic market-making has almost certainly led to lower spreads (for quantitative evidence consistent with this conclusion, see Brogaard 2010 and Hendershott et al. 2011). Indeed, it is hard to imagine traditional labour-intensive human market making, with its inevitably high costs and slow reaction times, being able to operate profitably at spreads of only one cent.

Because HFT, especially electronic market-making, contributes crucially to low spreads, and because low spreads are attractive to those who wish to trade securities, a symbiotic relationship between exchanges (or other trading venues with public limit-order books) and electronic market-making has come in to being. To reduce spreads and keep them low, trading venues need to provide the infrastructure (discussed in the next section) that makes electronic market-making possible. They also now almost all give ‘rebates’ to liquidity providers (those, such as electronic market-makers, whose systems have posted a limit order that is subsequently matched with a later order). The first trading venue to introduce the practice, in the late 1990s, seems to have been Island, which paid liquidity providers 0.1 cents per share, while charging those who ‘took liquidity’ (in other words, submitted the later order that was executed) 0.25 cents per share (Biais, Bisière and Spatt 2003, p. 6). Rebates form a major component of the revenue of electronic market-makers (although we know of no estimate of the exact proportion). Attracting electronic market-makers is a crucial component of the success of newly-created venues (see Jovanovic and Menkveld 2010 for the case of Chi-X and Dutch stocks), and they are also sometimes significant shareholders in those new venues. Indeed, one of the most important new venues, BATS, was created primarily by a team from Tradebot. Set up in Kansas City in 2006, BATS had by 2010 attracted around 10 percent of all US share trading ([www.batstrading.com](http://www.batstrading.com), accessed 30 March 2011).

## **2. Time, Space and the Materiality of Prices**

How is the speed essential to high-frequency trading achieved? It is common to think of modernity (or in some conceptualizations, postmodernity), as involving what Harvey (1989) calls 'time-space compression':

I use the word 'compression' because ... the history of capitalism has been characterized by speed-up in the pace of life, while so overcoming spatial barriers that the world sometimes seems to collapse inwards upon us. ...[S]pace appears to shrink to a 'global village'. (Harvey 1989, p. 240)

In high-frequency trading, this is only half right. Time shrinks, but space doesn't. Certainly, automated systems can respond to changes in order books far faster than human beings can. Human reaction times typically range from around 140 milliseconds for auditory stimuli to 200 milliseconds for visual stimuli (Kosinski 2010), while a state-of-the-art trading system's reaction time is well below a millisecond. That ultrafast reaction time, however, highlights how important spatial distance has become. Even at the speed of light in a vacuum, it would for example take a signal around four milliseconds to reach Nasdaq's matching engines (which are in a data centre in Carteret, NJ), from Chicago, which is around 800 miles distant. No matter how fast it could react, a system in Chicago would, therefore, be at a hopeless disadvantage compared to one closer to Carteret.

What such systems need to react to is a near-continuous stream of material representations of every change in order books: every new order, including its price and size, every cancelled order, and every executed trade (in the U.S., these 'raw' data feeds in aggregate now often exceed 3 million messages/second: see <http://www.marketdatapeaks.com>). To be able to receive these prices and other data as early as possible, and to have their reactions to it (their new orders and cancellations) reach the matching engines as quickly as possible, high-frequency trading firms rent space for their computer servers in the same building as an exchange's or other trading

venue's engines. 'Co-location', as this is called, is expensive and an important revenue source for trading venues (the rent for just one rack – a cabinet, usually around 2 metres high, that can hold multiple servers – can be as much as \$10,000 per month: Lex 2010) making co-location another manifestation of the economic symbiosis between venues and HFT.

Co-location sites are now often large buildings, and because of that they are not normally sited in traditional finance districts such as Manhattan, the Chicago Loop or City of London (with their high real-estate costs), but in cheaper, nearby places. The main share-trading matching engines in the U.S., for example, are all in data centres in northern New Jersey. For example, NY4, a large data centre across the Hudson River from Manhattan in Secaucus, houses matching engines for the share-trading venue Direct Edge, and also for EBS (the world's leading foreign-exchange trading platform) and the Boston Options Exchange. The New York Stock Exchange's engines are in Mahwah, Nasdaq's (as noted) in Carteret and BATS's in Weehawken. In early 2012 the engines of the most important derivatives exchange, the Chicago Mercantile Exchange, were relocated to a big new purpose-built data centre in Aurora, Illinois, forty miles west of downtown Chicago. (The situation in the UK is similar. While the London Stock Exchange's data centre is still in central London, most other venues have their matching engines in Slough or Basildon.)

These data centres are essentially large, frigidly-cold warehouses, consuming huge amounts of electric power: at the HFT conference mentioned in the introduction, it was reported that data centres have made the finance sector the largest power consumer in New Jersey. Dissipating the resultant heat is perhaps the biggest technical challenge in data-centre design, because heat is the most direct physical enemy of fast, reliable computing, and the necessary cooling systems are what makes data centres uncomfortably cold. They house, however, only small numbers of staff, who can spend most of their time in heated offices, controlling access, monitoring surveillance cameras, and being on hand to tackle technical malfunctions. Data centres are packed with row upon row of racks of computer servers and

digital switches. Miles of cabling – sometimes copper, but if the cable is more than around 100 metres long, fibre optics are faster – connect those servers to the matching engines and to the outside world. The servers owned by a trading firm will often be housed in a locked metal cage, to stop competitors or outsiders gaining unauthorised access to them.

It is possible to site data centres at some distance from the administrative headquarters of exchanges or other trading venues and from the offices of trading firms, because what matters to HFTs is the distance from their servers to the matching engines, not the precise locations of the latter. However, the big electronic market-makers all operate across multiple trading venues, and arbitrage between venues is important to other forms of HFT. This makes the fibre-optic links *between* data centres crucial. Indeed, these economic and technical interconnections between exchanges have turned U.S. markets into what is essentially a large, highly complex, spatially distributed and partially unified technical system.<sup>5</sup> (The situation in Europe lags the U.S., but is moving in the same direction.)

Because Chicago is the traditional primary site of derivatives trading, and New York of share trading, the fibre-optic links between Illinois (originally Chicago, but now also Aurora) and New York/northern New Jersey are the U.S. system's 'spinal cord'. Until August 2010, the best one-way transmission time available on those links was around eight milliseconds. However, high-frequency trader Daniel Spivey persuaded venture capitalist James Barksdale to fund the creation of a new cable, which unlike the old routes (which largely follow railway lines) is as direct as possible. The project – conducted largely in secret – cost around \$300 million. To speed construction, 125 teams worked in parallel, in places even creating what are essentially little tunnels through the rock of the Allegheny Mountains. The resultant link runs from downtown Chicago (the South Loop) to New York and to the site of the southernmost of the New Jersey matching engines, Carteret. Leasing 'dark

---

<sup>5</sup> Such interconnections have been present for three decades, and were made vividly manifest by the 1987 crash. Until the last decade, however, the interconnections typically 'passed through' human beings, such as Chicago pit traders, at some point.

fibre' in the cable (i.e. fibre for one's own private use) shaves around 1.3 milliseconds off the previously fastest one-way time, and this enables the link's owner, Spread Networks, to charge fees reported to be as much as ten times higher than those of older routes (Steiner 2010).

The considerable economic importance of a time advantage of little more than a thousandth of a second indicates a crucial issue. 'The idea ... that everyone can know at all times what the price is' (Carruthers and Stinchcombe 1999, p. 353) violates special relativity, which postulates that the maximum speed of a signal is the speed of light in a vacuum. In 1999, that was not yet a matter of importance; by 2012, it has become a crucial material constraint on the social structure of liquidity. The postulate could conceivably be wrong: in a widely reported recent elementary-particle experiment, neutrinos have been observed apparently travelling at very slightly in excess of the speed of light. But – barring the overthrow of central aspects of modern physics – traders at a large distance from matching engines are permanently doomed to learn 'what the price is' much more slowly than those who co-locate. The material assemblages that make possible today's liquid markets are in that sense Einsteinian, not Newtonian.

### **3. Materiality and Legitimacy**

High-frequency trading and the material assemblages that underpin it raise many fascinating issues. Some of these are technological. For example, the problem of the synchronization of clocks, which Galison (2003) argues was central to the emergence to the theory of relativity, is a major practical issue. Some are economic. For example, much high-frequency trading relies on short-term price prediction, prediction which the efficient market hypothesis of financial economics suggests is impossible. (Even electronic market-makers need to practise prediction in order to reduce the risks inherent in maintaining only tiny spreads between their bids and their offers.) The viability of HFT as a business therefore suggests that market efficiency breaks down at the ultra-short time periods of HFT, and that in respect to the latter the efficient market



hypothesis is too much of an abstraction from the material assemblages of trading.

Limited space, however, means that we can focus here only on one of the issues raised by HFT, and indeed only on one aspect of that issue. The issue is the fierce controversy that has erupted over its legitimacy, and the aspect is fairness. Until late 2008, HFT attracted little wider attention, even in the finance sector. However, in December 2008 Sal Arnuk and Joseph Saluzzi of the Chatham, NJ, institutional brokers, Themis Trading, posted to their website an attack on HFT as constituting 'toxic equity trading order flow' (Arnuk and Saluzzi 2008). Controversy then spread rapidly in the blogosphere, for example via Zero Hedge ([www.zerohedge.com](http://www.zerohedge.com)), and reached the pages of the *New York Times* on 24 July 2009, in an article that set the tone for much subsequent media commentary by describing HFT as 'a way for a handful of traders to master the stock market, peek at investors' orders and, critics say, even subtly manipulate share prices' (Duhigg 2009). By 2011, automated trading was even the topic of a thriller (Harris 2011).

The resulting debate has been sharp, indeed often vitriolic. The defenders of HFT amongst our interviewees almost all cited its role – documented in most studies of it by economists – in providing liquidity. As Joseph Mecane of NYSE Euronext told Duhigg (2009): 'Markets need liquidity, and high-frequency traders provide opportunities for other investors to buy and sell'. Critics of HFT contest what they believe to be over-narrow definitions of liquidity used by its defenders, in particular the bid-offer spread, and contend that the prices at which liquidity is provided are unfair, especially to institutional investors trying to buy or to sell large blocks of shares. Everyone focuses on bid-offer spreads, one critic told us in an interview, but HFTs post their prices in anticipation of order flow: 'They know it's coming'. Prices, notes Muniesa (2007, p. 377) 'can be regarded as fair or unfair, as accurate or inaccurate, as good or bad'. The critics of HFT challenge the 'quality' of prices even in apparently liquid markets, precisely as Muniesa suggests.

The controversy over HFT has many aspects. Market stability is one prominent one, especially after the wild fluctuations of prices in the U.S. in the afternoon of 6 May 2010 (the so-called 'flash crash'). Although it seems clear that the trigger was an execution algorithm, rather than a HFT program (CFTC/SEC 2010), many HFTs seem to have stopped trading in the face of extreme price movements – some suspecting technical faults, rather than catastrophic events – contributing to an evaporation of orders from public limit order books, with the orders that remained sometimes being filled at bizarre prices (a cent, or \$99,999.99, the smallest and largest prices that can be entered into the price fields of share-trading electronic order books).

Of all the aspects of the controversy, however, fairness has been most central. Its importance highlights the co-presence, even in financial markets, of multiple 'orders of worth', as Stark (2009) calls them. The most visible form of evaluation, economic evaluation (*how much is this security worth?*), coexists with what, in particular in the United States (less so in Europe, at least historically) has been a pervasive concern with fairness, especially in the stockmarket, in which many retail investors participate. It is important to the legitimacy of a market, especially in the U.S., that it be seen not just as 'efficient' and 'liquid', but also as 'fair'. For example, as Angel and McCabe (2010, p. 3) point out, 'the words "fair", "unfair" or "fairness" are mentioned 130 times in the Dodd-Frank Wall Street Reform and Consumer Protection Act'.

Issues of fairness arise within HFT itself. Within a data centre, different firms' servers are inevitably going to be located at different distances from the matching engines, and such is the concern within HFT over even tiny time lags that this is an issue of some sensitivity. Trading venues have responded by imposing equal cable lengths so that time delays are equal. The resultant coils of fibre-optic cable (technically unnecessary, but needed for fairness) are a physical reminder that we are dealing here with 'the creation and assemblages of Spacing(s)' and 'Timing(s)', not simply with 'a priori ... space and time' (Jones, McLean and Quattrone 2004, pp. 723-724).

No such simple physical answer, however, has been found for the wider questions of fairness that pervade the controversy over HFT. Arnuk and Saluzzi, for example, argue that ‘latency arbitrage’ (the advantages that high-frequency traders get from co-location and ‘raw’ data feeds)<sup>6</sup> ‘raises serious questions about the fairness and equal access of US equity markets which have made them the envy of the world’ (Arnuk and Saluzzi 2009, pp. 1). David Donovan of Sapiient (a global consultancy with expertise in derivatives trading technology) told the TABB Forum (<http://tabbforum.com>) on 8 August 2011 that ‘HFTs profit every day because they’re playing poker and can see everyone else’s hands’. Chicago-based data-feed analysts Nanex posted to their website (<http://nanex.net>) the comment: ‘On ... Aug 5, 2011, we processed 1 trillion bytes of data ... This is insane. ... It is noise, subterfuge, manipulation. ... HFT is sucking the life blood out of the markets: liquidity ... [A]t the core, [HFT] is pure manipulation’.

The most vigorous defender of HFT has been Manoj Narang, head of the Red Bank, NJ, high-frequency traders Tradeworx. He argues that ‘As in other industries, traders use technology to gain a *competitive* advantage, not an *unfair* advantage ... HFTs use technology to compete *with each other*, not with long term investors ... As in other industries, adoption of technology leads to compression of profit margins and reduced costs for consumers’ (Narang 2010, p. 3, emphases in original). He dismisses the hypothetical ‘example of how an HFT trading computer takes advantage of a typical institutional algo VWAP order’ advanced by Arnuk and Saluzzi (2009, p.2). ‘Many accounts of “predatory algos” are bogus’, argues Narang, including Arnuk and Saluzzi’s. Such accounts reflect either ‘a lack of knowledge of market microstructure, or simply cynical attempts to scare the public and policy makers’ (Narang 2010, p.15).

Narang, however, does *not* defend all facets of current institutional arrangements, and the issue on which he focuses in this respect is particularly

---

<sup>6</sup> ‘Raw’ data feeds come directly from (and are sold by) a single exchange and do not involve the aggregation and processing needed to produce the more commonly used, SEC-mandated, multi-exchange ‘consolidated tape’. They therefore reveal order-book changes more quickly.

relevant to our argument because it concerns an entanglement of materiality and issues of fairness deeper than unnecessary coils of cable. To explain it, let us return to the hypothetical order book in figure 1. Imagine that an investor wants to buy 300 shares on the venue in question, and is prepared to pay the offer price of \$21.00. Her order will be matched with the 200 shares on offer at that price, and one might expect that the remaining 100 shares would become a new bid to buy in the order book at \$21.00, an improvement on price on the existing bids.

In many cases, however, this will not happen, at least not immediately. The trading venue in question will often not allow the new bid to be posted in the order book, because offers to sell at \$21.00 will still apparently be present in other venues' matching engines, *even if the investor has also bought all the shares available at that price on all other venues.* (The combined effects of SEC regulations and arbitrage mean e.g. that the best offer prices on all U.S. share-trading public limit-order books are usually identical across venues.) The reason for delay is that each venue is required by SEC regulations to check that the new bid does not 'lock the market' by equalling the best offer price elsewhere. (A 'locked market' is one in which the best bid price is equal to the best offer price, and is prohibited by the SEC.) For Einsteinian reasons, however, venues cannot instantaneously ascertain the contents of other venues' order books, and indeed (unlike HFTs) they rely for this purpose not on fast, 'raw' data feeds from those venues but on the slower 'consolidated tape' (see note 6). So the unfilled remainder of the investor's order will not be posted as a new bid until the consolidated tape reports the removal of all \$21.00 offers from the other venues' order books.

A minority of traders can, however, circumvent this constraint by using an 'intermarket sweep order'. This is an order that carries a digital 'flag' indicating that the firm placing it has also sent orders that will remove from other venues' order books any orders incompatible with the new order being placed in the order book as soon as it is received. The firm placing the order thus takes upon itself (or, in practice, its computer systems) responsibility for compliance with regulatory requirements such as the ban on locked markets,

removing that responsibility from the trading venue. To be able to employ intermarket sweep order 'flags', however, one has either to be a registered broker-dealer (which carries substantial additional costs), or to persuade one's broker-dealer to allow one to use the flags. Because checking regulatory compliance is complex, and penalties are significant, broker-dealers will typically allow only particularly valued and sophisticated customers such as HFTs to do this.

The advantage that HFTs may get from the use of an intermarket sweep order is that – in the example discussed above – they will quickly learn from the fast 'raw' data feeds from the other trading venues that the \$21.00 offers have all been hit (matched with bids). Then, their computer systems anticipating a rising price, '[m]any such HFTs will rush to form the new ...bid [\$21.00 in this example], and will circumvent the [SEC] Order Protection Rule by sending ISO [intermarket sweep] orders' (Narang 2010, p.17). Their bids will thus be posted in order books prior to the new \$21.00 bid from the original investor, and will thus receive time priority (they will be the first to be filled). The issue is not trivial, Narang asserts: 'This results in tens of millions of dollars (conservatively) of extra trading costs for investors (and profits for HFTs)' (Narang 2010, p.17). Clearly, that is only a rough estimate, but use of intermarket sweep orders is widespread: figures from March 2010 for the twenty US stocks with the highest average dollar volume of trading indicate that in most cases between a third and a half of volume is the result of such orders (Narang 2010, p.18).

#### **4. Conclusion**

Readers with detailed familiarity with trading prior to automation will recognise the generic phenomenon of which Narang is suggesting an example: a systematic advantage that a particular category of market participant enjoys as the result of a 'technicality' of trading that is opaque to outsiders. The advantage may be small and even inadvertent – the SEC did not plan to give

HFTs this advantage, nor, as far as we are aware, did HFTs seek it – but may be significant in its aggregate consequences.

Because generic phenomena of this kind are familiar, even if their current material manifestations are not, such readers may be prompted to question the novelty of the new ‘social structure of liquidity’ described here. Carruthers and Stinchcombe (1999, p. 353) quite consciously posited an ideal type, an ‘elegant abstract stock-market structure’. However, almost all of the critiques of the legitimacy and fairness of high-frequency trading could have been (and sometimes were) directed at markets prior to automation. Market makers on trading floors enjoyed time-space advantages, measured not in the milliseconds or microseconds of automated trading but in full seconds or even minutes, and for that reason membership of an exchange was a valuable asset that could command often very high prices. The flash crash has led to widespread calls for obligations to provide liquidity to be imposed on electronic market-makers, and sometimes even on other types of HFTs, calls that often invoke, implicitly or explicitly, the past obligations of New York Stock Exchange specialists and Nasdaq dealers to provide continuous liquidity. The past, however, was not in reality a golden age of liquidity: such obligations were not successfully imposed *in extremis*. Notoriously, for example, many Nasdaq market-makers stopped answering their telephones during the 1987 stockmarket crash. Nor is opportunism – a frequent accusation against HFTs – new. Although opportunism on traditional trading floors was curbed by the informal norms documented by Abolafia (1996), there was certainly unscrupulous behaviour. For instance, traders in open-outcry pits would sometimes tacitly cooperate to ‘bounce’ prices around by entering into small trades with prices either above or below current levels, in the hope that these artificial changes in prices would trigger external customers’ ‘stop loss orders’, from which the pit traders could profit: in effect, they were practising what would now be called momentum ignition.

Perhaps, then, all that has changed is that the attention that HFT’s novelty generates, and the Internet-fuelled ease with which criticism of HFT circulates, mean that what would once have remained private disquiet about

traditional market making has become public denunciation of its electronic counterpart? With the age of large-scale automated trading still less than a decade old, there is no decisive, empirical counter to the objection that what we have documented in this article are old 'social' issues in new 'technical' guises. As researchers within science and technology studies (STS), however, we share a deep conviction that such a conclusion would be wrong.

Fundamental to STS is the view that phenomena can rarely be parsed cleanly into separate spheres of 'the technical' and 'the social'. Outside of the limited contexts within which human beings interact with their voices and naked bodies alone, objects and artefacts are implicated in all 'social' relations, and the development and implementation of technologies are always shaped profoundly by 'social' (economic, cultural, political, organisational, etc.) circumstances. It is most unlikely, we believe, that the huge changes in the material infrastructure of liquidity we have discussed can plausibly leave 'social' relations unaltered, or that the system we have sketched – with its chilly, computer-filled warehouses linked by fibre-optic cables carrying millions of messages a second as close as possible to the speed of light – will behave in the same way as the older, far slower, more 'human' structure of liquidity. One harbinger of the difficult issues that will need faced is that the recent UK Government Office for Science 'Foresight' investigation of computer trading felt forced to commission a review (Angel 2011) of 'The Impact of Special Relativity on Securities Regulation'. The review points to the generic relativistic constraint which, e.g., gives intermarket sweep orders their importance to the material practice of HFT: 'Because information can only travel at the speed of [light], the current state of the market will appear differently in different geographic locations' (Angel 2011, p.11).

That Einstein's theory of relativity might be relevant to the regulation of financial markets would have seemed bizarre in 1999, when Carruthers and Stinchcombe published their insightful analysis of the social structure of liquidity. In 2012, it is no longer bizarre. The material foundations of liquid markets have shifted, and the consequences are only beginning to unfold.

There can be few more important tasks for the social studies of finance than understanding that shift and its consequences, and we hope that this article has contributed in a preliminary and modest way to a task that (because of its complexity) must be a collective effort of our field.

## **ACKNOWLEDGEMENTS**

We are enormously grateful to Aaron Pitluck and two anonymous referees for the time they spent providing lengthy and insightful critiques of the first version of this paper. The main source of support for the interviews was a grant (RES-062-23-1958) from the UK Economic and Social Research Council, and MacKenzie's and Pardo-Guerra's work on automated trading is now being supported by a grant from the European Research Council (FP7-291733: EPIFM). MacKenzie would particularly like to thank Matthew Pullan for making possible his participation in High-Frequency Trading World, Europe, 2011. Responsibility for all errors remains our own.



## References

- ABOLAFIA, M. Y. (1996) *Making Markets: Opportunism and Restraint on Wall Street*, Harvard University Press, Cambridge, MA.
- ANGEL, J. J. & McCABE, D. (2010) *Fairness in financial markets: the case of high frequency trading* [Online]. Available at <http://ssrn.com/abstract=1737887>, accessed 22 July 2011.
- ANGEL, J. J. (2011) *The impact of special relativity on securities regulation* [Online]. Available at <http://www.bis.gov.uk/foresight>, accessed 11 November 2011.
- ANON. (2011) Some like it not, *The Economist*, 20 August [Online]. Available at <http://www.economist.com/node/21526387/print>, accessed 26 August 2011.
- ARNUK, S. L. & SALUZZI, J. (2008) *Toxic equity trading order flow on Wall Street: the real force behind the explosion in volume and volatility* [Online]. Available at <http://www.themistrading.com>, accessed 7 September 2011.
- ARNUK, S. & SALUZZI, J. (2009) *Latency arbitrage: the real power behind predatory high frequency trading* [Online]. Available at <http://www.themistrading.com>, accessed 24 August 2011.
- BEUNZA, D., MACKENZIE, D., MILLO, Y. & PARDO-GUERRA, J.P. (2011) *Impersonal efficiency and the dangers of a fully automated securities exchange* [Online]. Available at <http://www.bis.gov.uk/foresight>, accessed 11 November 2011.
- BEUNZA, D. & STARK, D. (2004) 'Tools of the trade: the socio-technology of arbitrage in a Wall Street trading room.' *Industrial and Corporate Change*, vol. 13, pp. 369-400.
- BIAIS, B., BISIÈRE, C. & SPATT, C. (2003) *Imperfect competition in financial markets: ISLAND vs NASDAQ* [Online]. Available at: <http://papers.ssrn.com/abstract=302398>, accessed 27 April 2011.
- BROGAARD, J. A. (2010) *High-frequency trading and its impact on market quality* [Online]. Available at <http://ssrn.com/abstract=1641387>, accessed 24 August 2011.
- CALLON, M., Ed. (1998) *The Laws of the Markets*, Blackwell, Oxford.
- CARRUTHERS, B. G. & STINCHCOMBE, A.L. (1999) 'The social structure of liquidity: flexibility, markets, and states.' *Theory and Society*, vol. 28, pp. 353-382.

- CFTC/SEC (2010) *Findings regarding the market events of May 6, 2010: report of the staffs of the CFTC and SEC to the joint advisory committee on emerging regulatory issues*. Washington, DC, Commodity Futures Trading Commission and Securities and Exchange Commission, 30 September.
- CHABOUD, A., et al. (2009) *Rise of the machines: algorithmic trading in the foreign exchange system*, Washington, DC: Board of Governors of the Federal Reserve System, International Finance Discussion Paper 980.
- DIMAGGIO, P. & POWELL, W.W. (1983) 'The iron cage revisited: institutional isomorphism and collective rationality in organizational fields.' *American Sociological Review*, vol. 48, pp. 147-160.
- DUHIGG, C. (2009) 'Stock traders find speed pays, in milliseconds.' *New York Times*, 24 July [Online]. Available at <http://global.factiva.com>, accessed 5 September 2011.
- DURBIN, M. (2010) *All About High-Frequency Trading*, McGraw-Hill, New York.
- GALISON, P. (2003) *Einstein's Clocks, Poincaré's Maps: Empires of Time*, Norton, New York.
- GETCO (2010) *A modern market maker's perspective on the European financial markets regulatory agenda* [Online]. Available at <http://www.getcollc.com>, accessed 26 August 2011.
- HARRIS, R. (2011) *The Fear Index*, Hutchinson, London.
- HARVEY, D. (1989) *The Condition of Postmodernity: An Enquiry into the Origins of Cultural Change*, Blackwell, Oxford.
- HASBROUCK, J. & SAAR, G. (2010) *Low-latency trading* [Online]. Available at <http://ssrn.com/abstract=1695460>, accessed 25 August 2011.
- HENDERSHOTT, T., JONES, C.M. & MENKVELD, A.J. (2011) 'Does algorithmic trading improve liquidity?' *Journal of Finance*, vol. 66, pp. 1-37.
- JONES, G., McLEAN, C. & QUATTRONE, P. (2004) 'Spacing and timing.' *Organization Studies*, vol. 11, pp. 723-741.
- JOVANOVIC, B. & MENKVELD, A.J. (2010) *Middlemen in limit-order markets* [Online]. Available at <http://ssrn.com/abstract=1624329>, accessed 4 April 2011.

- KOSINSKI, R. J. (2010) *A literature review on reaction time* [Online]. Available at <http://biology.clemson.edu/bpc/bp/Lab/110/reaction.htm>, accessed 28 January 2011.
- LEINWEBER, D. (2009) *Nerds on Wall Street: Math, Machines, and Wired Markets*, Wiley, Hoboken, NJ.
- LENGLET, M. (2011) 'Conflicting codes and codings: how algorithmic trading is reshaping financial regulation.' *Theory, Culture and Society*, vol. 28, pp. 44-66.
- 'LEX' (2010) 'Fast and furious.' *Financial Times* 15 January: p. 18
- MUNIESA, F. (2005) 'Contenir le marché: la transition de la criée à la cotation électronique à la Bourse de Paris.' *Sociologie du travail*, vol. 47, pp. 485-501.
- MUNIESA, F. (2007) 'Market technologies and the pragmatics of prices.' *Economy and Society*, vol. 36, pp. 377-395.
- NARANG, M. (2010) 'Tradeworx, Inc. public commentary on SEC market structure concept release' [Online]. Available at <http://www.tradeworx.com>, accessed 16 September 2011.
- O'BRIEN, R. (1992) *Global Financial Integration: The End of Geography*, Pinter, London.
- PARDO-GUERRA, J. P. (2010) 'Creating flows of interpersonal bits: the automation of the London Stock Exchange, c.1955-90.' *Economy and Society*, vol. 39, pp. 84-109.
- PITLUCK, A. Z. (2011) 'Distributed execution in illiquid times: an alternative explanation of trading in stock markets.' *Economy and Society*, vol. 40, pp. 26-55.
- PREDA, A. (2009) 'Brief encounters: calculation and the interaction order of anonymous electronic markets.' *Accounting, Organizations and Society*, vol. 34, pp. 675-693
- PRYKE, M. (2010) 'Money's eyes: the visual preparation of financial markets.' *Economy and Society*, vol. 39, pp. 427-459.
- SCOTT, S.V. & BARRETT, M.I. (2005) 'Strategic risk positioning as sensemaking in crisis: the adoption of electronic trading at the London International Financial Futures and Options Exchange.' *Strategic Information Systems*, vol. 14, pp. 45-68.
- SEC (2010) 'Concept release on equity market structure.' *Federal Register* 75/13 (21 January), pp. 3594-3614.

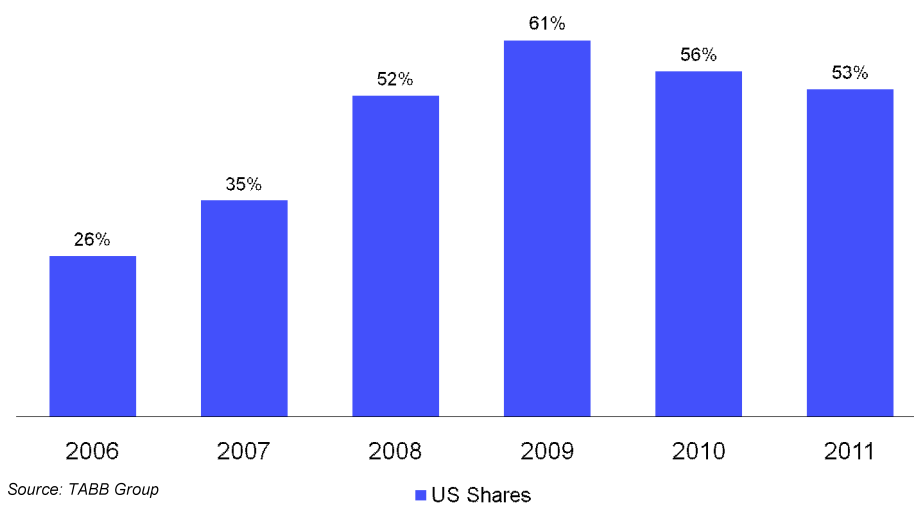
STARK, D. (2009) *The Sense of Dissonance: Accounts of Worth in Economic Life*, Princeton University Press, Princeton, NJ.

STEINER, C. (2010) *Wall Street's speed war* [Online]. 27 September .  
Available at <http://www.forbes.com>, accessed 30 August 2011.

ZALOOM, C. (2006) *Out of the Pits: Trading and Technology from Chicago to London*, Chicago University Press, Chicago.

'Bids' to buy		'Offers' to sell	
		↑	
		\$21.04	350
		\$21.03	400
		\$21.02	500
		\$21.01	150
		\$21.00	200
100	\$20.99		
450	\$20.98		
500	\$20.97		
100	\$20.96		
600	\$20.95		
	↓		

**Figure 1: A (hypothetical) limit order book**



**Figure 2: High-frequency trading's share of U.S. equity trading volumes.**

The TABB Group's counting rule (author interview) is e.g. that the HFT share is 25% if two shares traded as follows: non-HFT – non-HFT; HFT – non-HFT.

Electronic market-making (inc. equity trades prompted by market-making in derivatives)	42%
High-frequency statistical arbitrage	8%
Other HFT	6%
All HFT	56%

**Table 1: Shares of 2010 equity trading volume in U.S.**

Source: TABB Group response to IOSCO consultation, <http://tabbforum.com>, 10 Aug 2011.