A Sociology of Algorithms: High-Frequency Trading and the Shaping of Markets

Donald MacKenzie

June 2014

Second draft

Author's Address:
School of Social and Political Science
University of Edinburgh
Chrystal Macmillan Building
Edinburgh
EH8 9LD
Scotland
D.MacKenzie@ed.ac.uk

ABSTRACT

Computer algorithms are playing an ever more important role in financial markets. This paper proposes and exemplifies a sociology of algorithms that is (i) historical, in that it demonstrates path-dependence in the development of automated markets; (ii) ecological (in Abbott's sense), in that it shows how automated high-frequency trading (HFT) is both itself an ecology and also is shaped by other linked ecologies (especially those of trading venues and of regulation); and (iii) "Zelizerian," in that it highlights the importance of boundary work, especially of efforts to distinguish between (in effect) "good" and "bad" actors and algorithms. Empirically, the paper draws on interviews with 43 practitioners of HFT, and on a wider historical-sociology study (including interviews with a further 44 people) of the development of trading venues. The paper investigates the practices of HFT and analyses (in historical, ecological, and "Zelizerian" terms) how these differ in three different contexts (two types of share trading and foreign exchange).

INTRODUCTION

What becomes of economic sociology when markets and most participants in them are computer algorithms? That is now the case for many financial markets, such as in U.S. shares and U.S. and European futures. Analyzing price data from multiple exchanges for 2006-11, Johnson et al. (2012) argue that a transition has taken place from a "mixed human-machine" environment to an "all-machine ecology" in which "machines dictate price changes." They identify large numbers of very short time periods — many too short for human beings to intervene — in which prices crash or spike (by ± 0.8 percent or more) and then recover. A crash or spike that lasts, for example, only 25 milliseconds must, they infer, be machine-driven (Johnson et al. 2012, pp. 5 and 10).

Such claims are suggestive rather than definitive: as will be discussed below, there are difficulties in establishing even basic empirical facts such as the relative proportions of trading for which human beings and algorithms are responsible.[1] Nevertheless, there is persuasive evidence that some (but by no means all) financial markets have now moved into the third of three broad configurations:

1. Market actors are all human beings, and "the market" involves direct interaction among human beings.
2. The market is an algorithm (supply and demand meet in a computer system), but the actors remain mostly human beings; they interact with the market via computer screen, keyboard and mouse.

[1] The term "algorithm" is used in the sense in which interviewees use it, to refer not just to a set of instructions that is sufficiently precise to be turned into a computer program but to that program running on a physical machine and having effects on other systems.

3. The market is an algorithm, and most actors in it are also algorithms.

Nearly all existing sociological studies of electronic trading (e.g., Zaloom 2006; Preda 2009a and 2013; Saavedra, Hagerty, and Uzzi 2011) are of the second configuration or of the remaining human actors in the third configuration. The literature contains only glimpses of the third configuration: what Knorr Cetina (2013) calls "the interaction order of algorithms" remains largely opaque to economic sociology.[2] The most extensive — but still relatively brief — empirical discussion is Lenglet's (2011) ethnographically-based examination of the use in a brokerage firm of the "execution algorithms" discussed in the third section below, and of the resultant issues of regulatory compliance. There is a nascent sociological literature on algorithms more broadly (see, e.g., Mackenzie 2006; for an exemplary study of one particular algorithm, see Poon 2007 and 2009), but again there is a tendency to focus on algorithms with which human beings interact directly, such as the PageRank algorithm in Google (see, e.g., Hillis, Petit, and Jarrett, 2013).

Clearly, Latour and Callon's "actor-network theory" (see, e.g., Latour 2005) and Callon's actor-network economic sociology (e.g., Çalişkan and Callon 2009 and 2010) are pertinent when most market participants are algorithms. Actor-network theory is prepared to use the term "actor" to refer to non-human entities such as algorithms. While this usage remains controversial, it would plainly be a mistake to treat trading algorithms simply as the faithful delegates of human beings. As Adrian Mackenzie notes, "[a]n algorithm selects and reinforces one ordering at the expense of others" (2006, p. 44), but that ordering may not be the one its human

[2] Muniesa (2003, pp. 305-307) describes the use of algorithms to trade at the close of trading on the Paris Bourse; Beunza and Stark (2004, pp. 393-396) discuss the use of "robots" (statistical arbitrage programs that most likely implemented their trades automatically); Preda (2013, pp. 41-43) discusses human traders seeking to identify the traces of algorithms in market data.

programmers intended. Unexpected behavior by trading algorithms has led to well-publicized disasters, such as the $440 million loss incurred in 45 minutes by Knight Capital on August 1, 2012 when an old, forgotten algorithm mistakenly left on one of Knight's trading servers suddenly sprung to life.[3] Indeed, human users of algorithms may not always accurately understand even their routine behavior:

> [S]omeone could be in all honesty saying [their algorithms are] doing [something] when in fact they are doing something else: they're just not measuring it right. (Interviewee AP)

However, to develop a sociological analysis of automated trading it is necessary to go beyond generic actor-network considerations to more specific matters. Amongst the best sociological work on the relationship between "markets" and "technologies" is that of Knorr Cetina and Preda. From the viewpoint of this paper, their work can be read as examining different solutions to a generic issue facing all economic systems: how to coordinate economic action across space. They identify two broad solutions that correspond roughly to the first two of the three configurations sketched above:

1. A "network-based architecture" in which "coordination emerges from passing things through the pipes that link the network nodes" (Knorr Cetina and Preda 2007, p.116). Preda (2006), for example, has shown how one such "pipe" ─ the stock ticker, a telegraph-style device that

[3] A portion of the old software that kept track of the execution of orders had been moved, with the effect of making it non-operational. In consequence, when the old algorithm was inadvertently triggered on August 1, it kept sending out streams of orders, even though earlier orders had actually been executed, thus accumulating a giant unintended trading position. Unfortunately, Knight staff appear initially to have thought that the problem was likely to be the result of new, recently installed trading software. They therefore uninstalled it from Knight's other servers, which seems to have made things worse (Securities and Exchange Commission 2013).

relayed in close to real time the prices at which shares had been bought and sold and the number of shares transacted ─ reshaped stock markets.

2. A “flow” architecture or “scopic mode of coordination,” based on “collecting and ‘appresenting’ things simultaneously to a large audience of observers,” especially via computer screens, and “assembl[ing] on one surface dispersed and diverse activities, interpretations and representations which in turn orient and constrain the response of an audience” (Knorr Cetina and Preda 2007, pp. 116 and 126).

In the late twentieth century, scopic coordination became the dominant way financial-market activity was coordinated across space. However, as Knorr Cetina (2013) has suggested, the form of algorithmic trading on which I focus here — high-frequency trading or HFT — at least partially undermines that dominance. HFT operates at speeds beyond human capabilities. “If you’re sending [market data] to a human,” you have to slow it down, said an interviewee, because otherwise it becomes an uninterpretable blur on screen: “you can’t see it.” As will be noted below, HFT gives renewed significance to specific “pipes,” but does not simply reinstate an older architecture. Instead, examining the practices of HFT reveals a diversity of architectures in modern electronic markets, involving at least two quite different ways in which economic action is coordinated across space.

That much algorithmic action has effects that human beings “can’t see” has effects beyond those on coordination: it helps give rise to much suspicion of HFT. Michael Lewis’s *Flash Boys* (Lewis 2014), which portrays HFT in a most unflattering light, jumped almost immediately to the top of the *New York Times* non-fiction best-seller lists, and film rights in it were quickly bought by Columbia Pictures.[4] The popularity of a book that asserts that the U.S. stock market is "rigged" in favor of HFT indicates that HFT's legitimacy is precarious. Although the reasons for this are several (and not the focus of the research reported on here), one of them seems to be the undermining of one of the key ways financial markets have historically been legitimated. As stock markets were bounded off and framed as legitimate institutions, and as a boundary was drawn between "gambling" and "investment," finance was often portrayed as a domain in which "success should be ensured by constant and diligent observation" (Preda 2009b, p. 20).

So how can one legitimate a domain that sometimes seems no longer observable, at least not to those without specialist datafeeds and algorithmic equipment? The fear that HFT creates "a treacherous market ruled by machines" (Anon. 2010) is to be found not just among lay investors and the wider public but also within the financial markets themselves, and in some — but not all — of the markets discussed below provides a motivation for what, following Gieryn (1999), can be called "boundary work." As Zelizer (2012, p.145) observes:

> In all economic action … people engage in the process of differentiating meaningful social relations. For each distinct category of social relations, people erect a boundary, mark the boundary by means of names and practices, establish a set of distinctive understandings that operate within that boundary, designate certain sorts of economic

[4] Lewis focuses mainly on IEX, one of the trading venues that seek to exclude or place barriers in the way of algorithms deemed "opportunistic" or "predatory" (these venues are discussed below in the section "Dark Pools and Algorithmic Boundary Work"). His account of IEX is useful, and there is a delightfully vivid account of the construction of a new high-speed fiber-optic link between Chicago (where futures are traded) and northern New Jersey (where shares are traded), but the book as a whole offers only limited insights into the practices of HFT, which are viewed via the lens of HFT's opponents.

> transactions as appropriate for the relation, bar other transactions as inappropriate…

Some trading venues seek (in effect) to differentiate between “good” and “bad” algorithms, and how they do so resonates with efforts within HFT to delineate a sphere of unproblematically legitimate algorithmic action and separate it off from other forms of HFT. The consequences of ending up on the wrong side of the boundary can be sharp. Strikingly, even in some modern, competitive markets, firms whose trading algorithms make too much profit can find themselves expelled from trading venues as “predatory” or electronically stigmatized as “opportunist.”

Trading venues are the most immediate environment within which HFT algorithms act. It is difficult — though not entirely impossible — for them to act unless the consummation of deals takes place within a computer system. How that has happened historically has begun to be investigated by sociologists: see Muniesa (2003 and 2005) on the automation of the Paris Bourse, Pardo-Guerra (2010a and 2010b) on the London Stock Exchange, and Beunza and Millo (2013) on the New York Stock Exchange. The key argument in this literature is laid out most clearly by Muniesa (2011), who conceptualizes the mechanization of a market as a process of “explication” or (in the terminology of Deleuze 1990) of “expression.” Mechanization, Muniesa argues, is not “the laborious unveiling … of something that is already there, implicit,” not for example simply the direct translation of existing human processes into software. Instead, it is “a creative, performative, generative, provocative process” (Muniesa 2011, p. 2): there are different ways to turn a market into an algorithm, and the choices involved ─ including the apparently “technical” choices ─ are sometimes fiercely contested. They are also often highly consequential: the development of automated markets exhibits “sequence effects” (Abbott 2001, p. 286

and passim.) and path-dependencies (Arthur 1984, David 1992). Past choices ─ sometimes reflecting very specific, local priorities ─ facilitate and constrain present possibilities. This implies that an adequate sociology of HFT and other trading algorithms must be a historical sociology: it must examine not just current practices but the past choices and events that shape them.

That historical sociology turns out also to have to be a political sociology in the sense of Fligstein (1996 and 2001). "Markets are politics," as Fligstein argues. The sociotechnical "structures of markets" are indeed frequently "attempts to mitigate the effects of competition with other firms" (Fligstein 1996, p. 656), and incumbent market participants have typically either resisted the automation of financial markets and the emergence of HFT, or sought to shape automation so as to minimize the threat it poses. "[S]tates play an important role in the construction of market institutions," as Fligstein (1996, p. 600) notes: for example, as will be shown below, the operations of HFT algorithms trading U.S. shares are strongly shaped by the legacy of the efforts of a government regulatory body, the Securities and Exchange Commission (SEC), to reform share trading.

A natural framework for the necessary historical sociology of HFT is offered by Abbott's (2005) "linked ecologies."[5] An "ecology," in Abbott's sense, is a domain

[5] If this paper were a study simply of one domain, then conceptualization of that domain as a "field" would be preferable to conceiving it as an "ecology," given that the more extensive recent usage of "field" has provided rich intellectual resources (see, e.g., Fligstein and McAdam 2012). However, the focus here is on the relations among domains, and in that respect the model of "linked ecologies" is more flexible "topologically" than is often the case with fields (Abbott n.d.). That flexibility is necessary here for reasons discussed below, and Abbott's notion of "hinge" is particularly apposite. The "Chicago" heritage of the notion of "ecology" also renders the necessary "sociotechnical" extension of that notion rather more natural than might be the case with the notion of "field." For another use of "linked ecologies" to analyze finance, see du Gay, Millo, and Tuck (forthcoming).

"best understood in terms of interactions between multiple elements that are neither fully constrained nor fully independent." The term captures well the characteristics of HFT. In an ecology, "the elements constrain or contest each other," rather than behaving in an entirely atomistic way or their behavior being fully determined by a social structure or, one might add, by a technical system (Abbott 2005, p. 248). HFT algorithms interact directly with other algorithms and indirectly with each other, rather than acting in isolation or being parts of a unified technical system. HFT firms (which are typically small, privately-held proprietary trading firms, essentially combinations of people, significant but not huge amounts of capital, algorithms, and computer and communications hardware) indeed "constrain or contest each other," jossling for what Abbott calls "locations," for "things [they] are attempting to control" (Abbott 2005, p. 250): here, for market share and sometimes also for legitimacy.

Crucially, however, HFT is a *linked* ecology, one of a set of ecologies "each of which acts as a (flexible) surround for others" (Abbott 2005, p. 246). As already suggested, HFT is linked to the ecology of trading venues (and again "ecology" is an appropriate conceptualization: there are multiple trading venues in most of the main domains of automated trading, and they too compete for both market share and legitimacy) and to the ecology of regulation (also, at least in the U.S., a competitive sphere). None of these spheres entirely encloses the others without itself being enclosed by them. Regulation, say, might appear to be external to HFT and to trading venues, but it is not. For example, the single most powerful regulator of the trading of U.S. shares is the SEC, and in the 1975 Securities Acts Amendments it was tasked by Congress with linking U.S. share trading venues in such a way as to enhance competition. Tacitly, the SEC's goal was to undermine the dominant position of the New York Stock Exchange and its monopolistic "specialists" (market-makers). However, that goal was not achieved until — twenty years later — initially unrelated developments in trading venues (developments that were linked in their turn to the emergence of HFT) facilitated the task, and the SEC's achievement of its goal in its turn further encouraged HFT.

The development of HFT in U.S. shares was thus not the result of social-structural or technological determinism. Rather, it arose as an aspect of one of Abbott's "hinges": "strategies that work" in more than one ecology (Abbott 2005, p. 255). A "hinge" is not necessarily an alliance between actors in different ecologies. Here, it is a set of developments that, largely inadvertently, linked processes of change in different ecologies: HFT, trading venues, and regulation. The sedimented result of this historical process intimately shapes today's high-frequency trading of U.S. shares.

Two conceptual points about the invocation here of "linked ecologies" require clarification. First, following actor-network theory, this article makes no attempt to separate the "social" from the "technical." Each of the three ecologies discussed — HFT, trading venues, regulation — is a sociotechnical domain in which humans write algorithms and algorithms augment and diminish human capabilities, replace humans, and sometimes confound their plans. Second — and its flexibility in this respect is the key virtue of the idea of linked ecologies — the "topology" of how the three ecologies interacted was not historically fixed. HFT, for example, began as a "micro" activity largely enclosed in specific trading venues and shaped by more "macro" characteristics of those venues that were "social facts" that HFT's practitioners simply had to accept, but as HFT has developed it has come (in the case of U.S. shares, but not, e.g., foreign exchange) partially to enfold those venues: they are now shaped by it, more than vice versa.

In brief summary, then, this paper argues for and exemplifies a sociology of algorithms that is (i) historical (identifying, for example, path-dependencies); (ii) ecological (in Abbott's sense); and (iii) "Zelizerian" (examining boundary work and efforts to distinguish "good" from "bad" actors and algorithms). The paper proceeds as follows. After discussing the methods employed and explaining the overall forms of interaction among the three main types of algorithm discussed here (trading venues' matching engines, which consummate trades; execution algorithms used by institutional investors to buy or sell large blocks of shares; and HFT algorithms), the paper then focuses in more detail on how HFT algorithms act. Next, the historical process sketched above — the emergence of a "hinge" connecting HFT in U.S. shares, developments in trading venues, and the regulatory ambitions of the SEC — is described in more detail. Finally, the differences between the practices of HFT in three different markets (one main case, and two comparator cases) are examined to show how historical processes, different linked ecologies, and boundary work shape those practices:

i. U.S. "lit" trading venues for shares.[6] This is the main case discussed: these venues and trading on them are directly shaped by two factors that are in tension: first, the "hinge" connecting the development of trading venues to that of HFT and to regulation; and second, path-dependence, in the form of the legacy of a little-known late-1970s' decision in this domain about how to coordinate economic action across space. The result is a deep contradiction in that coordination, a contradiction that, *inter alia*, has made a specific form of algorithmic

[6] As discussed below, a "lit" venue is one in which the electronic order book is visible to the humans and algorithms that trade on the venue; in a "dark" venue it is not visible.

action — the Intermarket Sweep Order — pivotal to the high-frequency trading of U.S. shares.

ii. U.S. “dark” trading venues for shares. These are markets in which the boundary work outlined above (drawing what are sometimes seen as “moral” distinctions among algorithms) is prominent. Here, competition among venues for market share and competition for legitimacy are interwoven intimately, and among the consequences are that the behavior of HFT algorithms is subject to direct surveillance by venues, with some HFT firms excluded or electronically stigmatized as “opportunistic.”

iii. Foreign-exchange trading venues. These markets, vitally important in themselves, also serve as a historical comparator case to U.S. share trading. With no equivalents of the SEC and the historical “hinge,” and a different linked ecology, boundary work takes a simpler, more brutal form, and how economic action is coordinated across space is also different from in share trading. In foreign exchange too that coordination is marked by a contradiction, but a different contradiction that gives salience to a different algorithmic action: “last look.”

METHODS

High-frequency trading is a difficult domain to research either quantitatively or qualitatively. The barrier to quantitative research is that, with very limited exceptions, financial-market data available to researchers do not contain datafields that indicate whether the participants in a transaction were humans or algorithms, or if the latter whether the algorithm was a HFT algorithm.[7] In one important U.S. case, economists have gained access to data containing anonymized trading-account identifiers, in which HFTs are identifiable via their distinctive trading styles, but access to that Commodity Futures Trading Commission (CFTC) dataset is no longer available and publications based on it have been suspended.[8] In consequence, although two market consultancy groups (Tabb and Aite) publish figures on the proportion of trading that is HFT (see, e.g., Table 1), these figures are simply estimates that draw on what consultants' contacts tell them.

Qualitative research on HFT is also hard. HFT firms are, as noted above, most commonly privately-held partnerships that do not report publicly on their activities, and often go to some lengths to protect the confidentiality of their trading: as a former specialist in automated trading puts it, "[i]n this business, everyone knows that loose lips get pink slips" (Durbin 2010, p. 2). Despite this obstacle, however, 43 founders, employees, or ex-employees of HFT firms agreed to be interviewed by the author about the practices of HFT, the contingencies that bear upon these practices, and (in the case of interviewees with long experience of the sector) the history of HFT. (In the quotations from these interviews, interviewees are

---

[7] NASDAQ has made available to a number of financial economists a dataset for a sample of 120 stocks for 2008, 2009 and one week of February 2010. The dataset contains a field for each transaction that NASDAQ has populated, based on its informal knowledge of firms' business models, as HH, NN, HN, or NH. A transaction labelled HN, for example, is one in which a high-frequency trading firm ("H") hits a bid or lifts an offer posted by a non-HFT ("N"). The resultant work (especially Brogaard, Hendershott, and Riordan 2013 and Hirschey 2011) forms a useful crosscheck of some of this paper's interview-based findings.

[8] This dataset contained futures market data held by the regulator, the CFTC, on which the CFTC's Office of the Chief Economist (OCE) and some academic economists linked to the OCE had started to work. An early paper based on this dataset (Baron, Brogaard, and Kirilenko 2012) was reported by the *New York Times* on December 3, 2012, under the heading "High-Speed Traders Profit at Expense of Ordinary Investors" (Popper and Leonard 2012). Following a complaint from the Chicago Mercantile Exchange (Lewis 2014, p. 207), the CFTC suspended the publication of results from the analysis of this dataset, and to date it has not resumed.

identified chronologically from AA, the first high-frequency trader to be interviewed, in October 2010, to BQ, interviewed in April 2014. In the seven cases in which the same high-frequency trader was interviewed more than once, a numeral identifies which of the interviews is being cited.)

The interview sample was constructed in a variety of ways. One was by identifying, from published sources such as reports in the specialist press, as many as possible of the HFT firms active in Chicago, New York, London, and Amsterdam (the four most important sites of HFT worldwide). If those press reports or the firms' websites (not all HFT firms have publicly visible sites) identified the firms' founders or a named individual with responsibility for trading or technology, and if a telephone number could be found (some HFT firms' websites do not disclose their addresses or telephone numbers, but these can sometimes be obtained by other means) at least one of those individuals in each firm was then telephoned. This "cold calling" was successful in just over half the cases in which it was attempted, generating 15 interviews. Other ways of identifying interviewees were more ad hoc: approaching speakers at HFT industry meetings, using a list of potential interviewees provided by an industry analyst, snowballing from earlier interviewees, and happenstance contacts.[9]

Clearly, no claim of representativeness can be made for this sample, which is, for example, made up disproportionately of senior high-frequency traders: a happenstance contact and some of those identified via snowballing were quite junior, but most of the ways of identifying interviewees led to better known — and therefore more senior — people. Nor was it possible to follow even a semi-structured

---

[9] For example, a member of the audience at an early talk given by the author on HFT identified himself as a high-frequency trader.

interview schedule: the overwhelming need was to keep the conversation going, and keep it focused on the practices of HFT (the author struggled in one interview with two former Chicago pit traders who had joined an HFT firm to stop them simply talking about the pits). It was easy inadvertently to ask a question that disrupted the interview because the interviewee felt unable to answer it:

> Interviewee AD1: Some companies don't wait for the exchange to tell them what's trading.
>
> Author: Oh, so how do you manage to…?
>
> Interviewee AD1: That I can't… I mean not only would I lose my job, I might lose my legs too!

> Author: Do you use ISO [Intermarket Sweep Order] orders?
> Interviewee AF: Can't say.

However, information proffered by the early interviewees made it possible gradually to identify a set of HFT techniques that are widely known in the sector, widely practiced, and thus acceptable topics of questioning. Interviewees would say of such techniques: "everyone knows that" (interviewee AA); "today that's High Frequency Trading 101" (AH). Similarly, early interviews provided glimpses of contingencies affecting the practice of HFT in particular domains, contingencies (such as "last look" in foreign exchange) that were unknown to the author at the start of the research, but common knowledge in those domains. Again, it proved possible gradually to build up a sense of what those contingencies are, and in later interviews to focus more directly on them. Although the research was not originally designed as comparative, it became clear as the interviews proceeded that there were marked differences between the practices of HFT in different domains (especially between shares and foreign exchange), and later interviews focused in part on those differences.[10]

Because of these iterative aspects, the interviews with high-frequency traders were more like solving a jigsaw puzzle (with no picture to guide one) than conducting a survey. Fortunately, however, matters were more straightforward when researching the ecologies surrounding HFT. Interviews, for example, were easier to secure. (Including the HFT interviews, an overall total of 125 interviews were conducted: see Table 2.) Particularly in the case of U.S. shares, the overall histories of both the main trading venues (the New York Stock Exchange and NASDAQ) and of regulation are reasonably well documented, making it possible to draw on documentary sources for the broad picture, and to focus interviews with trading-venue personnel (44 people in total: see Table 2) on venues of three kinds: those that documentary sources and initial interviews suggested were pivotal in the development of HFT; those that set out to monitor the behavior of HFT algorithms and engage explicitly in boundary work; and foreign-exchange venues, because these offer an interesting comparator case to share-trading venues. When an interview in the wider set (i.e., beyond HFTs) is quoted below, the source is identified simply as "interviewee." To keep the number of interviews from becoming too large, it was decided not to interview regulators, because documentary sources on the development of regulation were adequate for the purposes of this paper.

[10] For reasons of space, the paper does not discuss the two remaining main domains of HFT, fixed income and futures. Although incumbents' resistance to automation has been more successful in fixed income than in foreign exchange (see Table 1), the contingencies shaping electronic trading in those two domains are broadly similar. Resistance in futures, however, collapsed, just as it did in shares, but via a process in which regulation played no direct part, and for contingent reasons (discussed in author ref.) U.S. futures trading remains dominated by a single venue, the Chicago Mercantile Exchange. It thus resembles what share trading might have looked like had the single, national Composite Limit-Order Book, discussed below, triumphed.

Despite the limitations of the financial-economics literature on HFT caused by the data problems referred to above, there is one crucial issue on which interview-based conclusions can be cross-checked against that literature: the capacity of HFT techniques to predict short-term price changes. This issue particularly needs checked, because that predictive capacity seems to fly in the face of the "efficient market hypothesis" of financial economics, which decrees price changes not to be predictable. Fortunately, as will be seen below, economists' quantitative findings support the interview-based conclusions of this paper in this respect.

MATCHING ENGINES, EXECUTION ALGORITHMS AND HFT ALGORITHMS

To say of a market that it is an algorithm is, in most of the markets discussed in this paper, to say that deals on it are consummated by a computerized matching engine that manages an electronic order book. To explain what a matching engine does, it is easiest to use a visual representation of an order book of the kind sometimes synthesized by computer systems for the benefit of the remaining human beings interacting with a market. Figure 1 is a screen shown to me by an interviewee testing one of the execution algorithms discussed below. It shows the order books on a number of trading venues for the shares of the New York savings and loan, Astoria Financial. (My interviewee was not aware that his algorithm was trading Astoria shares: when I asked him what the symbol "AF" stood for, he did not know.) On the left of the screen are bids to buy Astoria shares: for example, a bid or bids on NASDAQ to buy 192 shares at $7.74; a bid or bids on Arca to buy 800 shares, also at $7.74; and so on. On the right are the corresponding offers to sell.

Consider one of the venues trading Astoria shares, for example NASDAQ (see Figure 2). The crucial functions of a matching engine are to maintain a trading venue's order book and to search for bids and offers that match. In the book in Figure 2, there is no match. However, a match would be created immediately by a human or algorithm entering a bid to buy shares at $7.75 or below or an offer to sell them at $7.74 or above (an order that can be executed immediately is called a "marketable" order). Once the matching engine finds a match, it consummates the trade and sends the two parties electronic confirmations. Unlike in a traditional "human" market, no negotiation is involved; indeed, in share trading (but not always in foreign exchange) the whole procedure is entirely anonymous.

In addition to the remaining human participants in markets, two broad categories of trading algorithm interact with matching engines. The first is execution algorithms. These are used by institutional investors, or brokers acting on their behalf, to buy or sell large blocks of shares or other financial instruments. Execution algorithms break up those blocks and bring them to market in a way designed to minimize "market impact" (a large buy order, for example, will typically cause prices to rise before it is fully executed). For example, one standard class of execution algorithm is "volume participation" algorithms, which keep track of the volume of transactions over a rolling time period (a minute, for example), and place new orders that are a set proportion of that volume, the rationale being that market impact is typically lower when markets are active. The other broad category of trading algorithm is proprietary trading algorithms, of which the subclass on which this paper focuses is HFT algorithms.[11] Unlike an execution algorithm, a proprietary trading algorithm does not set out to buy or sell a specific quantity of the instrument being traded; indeed, HFT algorithms are almost always programmed with the goal of making a profit while not accumulating the risky trading position that would be created by buying a lot more than they sell, or vice versa.

By placing or cancelling orders, HFT and execution algorithms interact directly with matching engines, and via the latter interact indirectly with each other. Those who write execution algorithms design them to hide their activities from human professional traders, proprietary-trading algorithms, and even other execution algorithms. (Many execution algorithms are now just as sophisticated as most HFT algorithms, and employ similar techniques of price prediction.) As an interviewee put it, execution algorithms “take [a] huge order and chop it up into little tiny pieces and, if we do it right, anyone who’s looking at it can’t tell that there’s a big buyer: it looks like tiny, little retailish trades [i.e. trades by lay investors] … and no-one knows who or what is happening.” The reason for taking this approach is clear: if a proprietary algorithm can successfully detect the digital footprint of an execution algorithm that is (for example) in the process of buying a large block of shares, it can make money at its expense by buying shares ahead of it and selling them to it at a profit.

[11] Another type of proprietary trading algorithm is statistical arbitrage algorithms. Like most HFT algorithms, these also seek to predict patterns of price changes, but over a longer timescale — from a few minutes to several weeks or even months — and they typically employ different sources of prediction, seeking to identify factors, involving for example firms’ balance sheets or correlations among very large baskets of stocks, that shape stock price dynamics over these longer timescales. The boundary between HFT and statistical arbitrage is, however, fuzzy.

Patterns of algorithmic behavior can emerge that can indeed be understood only "in term of interactions between multiple elements that are neither fully constrained nor fully independent" (Abbott, as quoted above). For instance, two or more volume participation algorithms can start to influence each other's behavior. As interviewee AE put it, "every time one of them prints," in other words executes a trade, it boosts the volume of transactions, leading the others to seek to trade as well:

> It causes all the other guys [algorithms] to print, which causes the first one to print, and the stock will just go "zhwoom" [rise sharply] until they're all done [have made the programmed purchases] and then it'll go "pfft" [fall sharply] again.

(Interactions of this generic kind among algorithms are the most likely cause of the short-lived price spikes and crashes observed by Johnson et al. 2012.) Other algorithms programmed to spot episodes of "price momentum" can profit from the episodes described by AE, and there is even a further level of interactive behavior, AE reported. In this, the process of mutual influence among volume participation algorithms is deliberately simulated, with the goal of exploiting the "momentum" algorithms that "discover" such episodes.

THE PRACTICES OF HIGH-FREQUENCY TRADING

It would be quite mistaken, however, to imagine that the behavior of HFT algorithms is always, or even mostly, sophisticated, reflexive "gaming" of this kind. As the term "high-frequency" suggests, HFT is based on large volumes of trading, and intricate "gaming" strategies are unlikely to scale up successfully. Sometimes, HFT is as simple as detecting a financial instrument on sale on one venue at a lower price than it is being bid for on another, but such "arbitrage opportunities" (as market participants call them) are now small enough and infrequent enough, interviewees reported, that they too could not form the basis for a large-scale business. Rather, the core HFT practices are a variety of broadly applicable techniques of very short-term price prediction.

To give a flavor of those techniques, consider two of them that are used by the algorithms of the firms of all the HFT interviewees who were prepared to discuss such matters in any detail. The first is "order-book dynamics" for the instrument being traded. At its simplest, said interviewee AH, that is a matter of an algorithm calculating whether "the bid [is] bigger than the offer." Consider, for example, the order book in Figure 2, in which the best (i.e. highest) bid consists of 192 shares and the best (i.e. lowest) offer consists of 488, and imagine that this is the only place the shares in question are traded. The best offer is bigger than the best bid, suggesting greater immediate selling interest than buying interest and that, "probabilistically, the next [price] tick is likely to be [down]" (interviewee AF). (That this form of prediction works, and that HFT firms employ it, is one of the issues on which the interviews are supported by the financial-economics literature.)[12]

Another informative aspect of order-book dynamics is what interviewee AN called "time and sales": the transactions involving the instrument in question, and when and at what prices these took place. Note too that in the cases both of U.S. shares and foreign exchange, both of which are traded on multiple venues, a

---

[12] Using the NASDAQ dataset described in note 7, Brogaard, Hendershott, and Riordan (2013) show that the relative size of the best bid and offer does have predictive power, and the direction of HFTs' marketable orders is consistent with their trading being informed by this.

sophisticated HFT firm will take into account the dynamics of *all* the different order books for the instrument being traded.[13] In addition, HFT algorithms can also include, in their calculation of what interviewee AN called "book pressure," the sizes of bids below the best bid and offers above the best offer. However, this makes an algorithm's calculation more vulnerable to "spoofing," in other words to other algorithms or human traders placing bids or offers not with the intention of buying or selling but simply to create the impression of excess demand or supply. (In Figure 2, for instance, there are large bids at $7.72 and $7.71, but unless prices fell very fast they could be cancelled before being executed.) Spoofing is "a big problem," said interviewee BH, but there are ways of HFT algorithms defending themselves, reported interviewee AY, such as omitting or underweighting very recently placed bids and offers when the algorithm calculates the balance between the two, the rationale being that orders that have been in the book for longer are less likely to have been placed by a spoofer.

A second widely used HFT predictive technique involves the order books and price movements of financial instruments that are correlated with the instrument being traded, especially when those instruments are known typically to "lead" the latter. The most important single example is "futures lag": the use by HFT algorithms trading shares or exchange-traded funds of movements in the prices of, or changes in the order book for, the corresponding share-index futures.[14] The interviews — and, once again, the literature of financial economics (Hasbrouck 2003) — indicate that

[13] Lewis (2014) emphasizes that in the case of share trading (he does not discuss foreign exchange) either incentives to brokers or different speeds of connections to different venues can lead execution algorithms typically to bring orders to a particular favored venue first and to other venues only after small but consequential time delays.

[14] An index future is a derivative whose pay-off depends on the movement of an underlying share-price index such as the S&P 500 or NASDAQ 100; an exchange-traded fund is a stock whose price similarly tracks the aggregate prices of the shares making up such an index.

changes in index-future prices tend to lead those both of the corresponding exchange-traded funds and of the underlying shares. This makes index-future price and order-book changes a crucial predictor of changes in the prices of those funds and shares.

Order-book dynamics and prediction using correlated instruments are pervasive techniques, used, as noted, by the firms of all HFT interviewees prepared to discuss these matters. A third, more specialized source of information — by no means all firms used it — is macroeconomic or company-specific "news," which now is often disseminated in machine-readable form: an algorithm that can act on such news before it is fully incorporated into prices can profit handsomely. Different sources of information are not usually processed in isolation. HFT firms' algorithms typically aggregate multiple sources — order-book dynamics, data on multiple correlated instruments, perhaps news — in real time. Very commonly, but not universally,[15] the result of the aggregation is an automated estimate of the "theoretical value" (interviewee AG2), "fair value" (AG2 and AR), "theoretical price" (AO), "fair price" (AF), "perfect price" (AN) or "microprice" (AO) of the shares or other instrument being traded. These terms are synonyms; in the context of HFT, they mean "the price you can reasonably expect to transact at in the near future" (AG2), where the "near future" might be anything from less than a second to a couple of minutes. This price is most easily thought of as the dependent variable in a multiple regression, in which the independent variables are predictors such as the bid:offer imbalance, the prices of related instruments, etc. (BF). However, other firms' algorithms employed different forms of aggregation. Thus interviewee AN described

[15] Thus interviewees AI, AQ, AU, and BP reported that their algorithms' predictive efforts did not take the form of estimating a theoretical value.

an elaborate automated "polling" system in which the weights given to the "votes" of different predictors varied according to market conditions.

In the case of U.S. shares, the difference between this "theoretical value" and market price will often — indeed usually — be less than a cent. If his firm's algorithms "think the price should be [$]2.396," and there are bids for the shares at $2.40, they "will sell that security" at $2.40, said interviewee BJ. "Sometimes I will lose money on that trade … on average I will win maybe 55 percent of the time, which is enough for me." There is, however, a critical choice involved in exactly how HFT algorithms make these sales or purchases. They can act "passively" or "make liquidity" (as market participants put it): they can place in the electronic order book bids and/or offers with prices that cannot be executed immediately. Alternatively, they can act "aggressively" or "take liquidity": they can submit a marketable order that will be executed as soon as it is received by the matching engine. For example, in the book shown in Figure 2, an offer to sell shares at $7.74 is marketable, "aggressive," and "liquidity-taking"; an offer to sell shares at $7.75 is non-marketable, "passive," and "liquidity-making."[16]

As that higher price shows, liquidity-making has potential economic advantages. Other things being equal, a non-marketable order that another algorithm or human being later transacts against is executed at a more favorable price than a liquidity-taking order, and (at least in share trading in the U.S. and much of Europe) it also receives a "rebate": to encourage liquidity-making, the trading venue will make a small payment (around 0.3 cents/share) to a firm that has entered a liquidity-making

[16] In what follows, the more technical term "liquidity-making" is preferred to the more colloquial "passive," because in the context of HFT the connotations of the latter are misleading: algorithms that place "passive" orders are frequently frenetically active (see below).

order that has been executed against. Furthermore, the placement of liquidity-making orders inherits, in some contexts, the legitimacy of a traditional human role, that of the market-maker who always stands ready both to buy and to sell the instrument being traded. An important subcategory of HFT firms (represented in my sample by interviewees AC, AG, AO, AQ, AW, BG, BH, BI, BJ, BK and to some extent BE and BF) position themselves primarily as "electronic market-makers." In some contexts, indeed, the distinction between liquidity-making and "aggressive" liquidity-taking is freighted with moral significance. One interviewee, who was trying to persuade others in his automated but not fully high-frequency trading firm to shift their emphasis from making to taking liquidity, reported that their reaction was as if he had asked them "to stab their sister." (Such connotations are discussed below in the section "Dark Pools and Algorithmic Boundary Work.")

The primary activity of a market-making HFT algorithm is to keep its buy orders at or close to the best bid price ($7.74 in Figure 2) and sell orders at or close to the best offer price ($7.75) — its exact choices of prices and order sizes will depend on its calculations of theoretical value and the need for it to manage potentially risky holdings of the instrument being traded[17] — with the goal of having others execute against both its bids and its offers. The algorithm thus aims ideally to earn the "spread" between the bid and offer prices (most commonly one cent, as in this example), along with two rebates; i.e., a total of around 1.6 cents per share bought and sold. Market-making sounds simple, but isn't. A market-making algorithm needs constantly to place new orders and cancel existing orders as prices move, and its need to predict price movements is no less than that of an "aggressive"

[17] A market-making algorithm will often "shade" its prices for inventory reasons, for example posting particularly attractive offers to reduce its inventory of the asset being traded if that inventory has become uncomfortably large.

algorithm. If, for example, a market-making algorithm is making markets in the QQQs (an exchange-traded fund that tracks the value of the shares in the NASDAQ-100 index), and the price of NASDAQ-100 futures goes up, the market-making algorithm's offer prices almost instantly become "stale," and can profitably be "picked off" by an aggressive algorithm. So the market-making algorithm must cancel those existing offers as quickly as possible and replace them with offers at a higher price before (in market-making terminology) it is "run over."[18]

The risk of being "picked off" or "run over" is only one of the disadvantages of liquidity-making algorithmic action. There is no certainty when — or indeed whether — a non-marketable bid or offer placed in the order book will be executed (and, of course, if it is not executed the algorithm will never earn the spread or a rebate). In contrast, aggressive, liquidity-taking algorithmic action, employing marketable (and thus immediately executable) orders, offers much greater certainty. Under some circumstances, that certainty outweighs the economic disadvantage of those orders (that they involve "paying the spread" and earn no rebate), so long as the algorithm has identified a potential profit opportunity larger than the additional cost of aggressive action. There is also a certain cognitive advantage to aggressive action. HFT firms nearly always "back test" new algorithms extensively, simulating their performance using past market data. Aggressive, liquidity-taking algorithmic action "is an easier thing to simulate," said interviewee AY: with passive, liquidity-making

[18] Some of the algorithmic behavior condemned as predatory in Lewis (2014) could equally be explained by market-making algorithms avoiding being "run over." The execution of a large buy order, for example, will generally drive prices up, and — in part for the very reasons explored by Lewis — that coming rise will often be predictable from its traces in order-book dynamics. So market-making algorithms "realize" their offers have become stale (and will be loss-making if left in order books), and hence cancel them and replace them with offers at higher prices.

action you have a more difficult job "predict[ing] whether you would have gotten the fill or not."

There is a substantial degree of differentiation among HFT algorithms in respect to the actions they take: some predominantly make liquidity; some nearly always act "aggressively" and take liquidity. Indeed, that differentiation extends to the firms that employ them, which often seem largely to specialize either in liquidity-making or liquidity-taking. "[I]t's funny how there are some firms today who almost exclusively provide liquidity and other firms who almost exclusively take liquidity," said interviewee BE: "it's almost like two very different strategies and thought processes." (This is another point on which the interviews can be checked against a financial-economics study. A paper based on the Commodity Futures Trading Commission dataset that did temporarily enter the public domain [Baron, Brogaard, and Kirilenko 2012] found that "the aggressiveness of a given HFT firm [the degree to which its trading is liquidity-taking] is highly persistent" [p. 27].)

Despite the ethos of secrecy that surrounds at least some HFT firms, the interviews suggested that because of factors such as the movement of personnel between firms, the main techniques of HFT are common knowledge in the sector. "There are secrets but there are no secrets," was how interviewee AI put it. With HFT algorithms therefore often using similar predictive techniques, competition amongst them often boils down to relative speed.[19] To receive data on order-book charges with minimum delay, and to submit orders and cancellations of orders as quickly as possible, HFTs pay trading venues hefty fees to "co-locate": to place the servers on which their algorithms run in the same building as the servers on which the venue's

[19] Absolute speed matters too, for example in helping a market-making algorithm minimize the risks it is taking by adjusting its bids and offers as quickly as possible as market conditions change.

matching engines run.[20] U.S. share-trading HFTs that use futures prices as predictors have to invest in a particular "pipe," the faster possible links — four years ago, a new fiber-optic cable following a new, more direct route; now, a series of microwave towers — between Chicago (the main U.S. futures-trading matching engine is located in Chicago's outer suburbs) and the data centers in northern New Jersey in which shares are traded (see, e.g., Lewis 2014).

The importance of relative speed gives HFT something of the character of a technological "arms race."[21] It also gives salience to very specific features of matching engines, of the physical machines on which those engines run, and of the "order gateways," the trading-venue servers that process incoming orders before passing them to the matching engines and transmit "confirms" (messages to the computer systems of the parties to a trade telling them that one of their orders has been executed). Recall interviewee AD's fear that giving me a specific piece of information might cause him to "lose my job" and perhaps "my legs too." Two years after that October 2011 interview, it suddenly became clear what that piece of information was, when the *Wall Street Journal* (Patterson, Strasburg, and Pleven 2013) revealed that the order gateways of the Chicago Mercantile Exchange (the prime U.S. futures-trading venue) typically sent "confirms" one to ten milliseconds before news of the trade was disseminated on the exchange's wider datafeed.[22] (Interviewee AD later told me this was indeed what he had been unable to say.) That time difference is economically important. Consider, for example, an HFT employing

[20] Interviewees also reported that in some but not all data centers it is also possible to pay more to be in a part of the building close to the matching engines.
[21] For a formal model of HFT's "arms race" component, see Budish, Cramton, and Shim (2013).
[22] A trading venue's datafeed involves the aggregation of messages from all its matching engines, which can make it slower than an individual signal from one of the latter sent out via a particular order gateway.

“futures lag” to predict movements in the prices of shares or exchange-traded funds. If it was also making liquidity in those futures, it could “take [its own] fill as market data,” as another interviewee put it: when it received a “confirm” that its futures bids had been hit or its offers lifted, it could infer that prices were moving before those movements were apparent on the Chicago Mercantile Exchange datafeed.

Press reporting of such matters (and the anti-HFT account in Lewis 2014) can, however, implicitly give a misleading impression of the typical size of HFTs’ profits. Much of the high-frequency trading of shares, for instance, involves predicting a “tick” of prices up or down, and the unit of price for U.S. shares costing $1 or above is a cent. That latter figure gives a sense of the scale of routine potential profit: interviewee AF, for example, reported that for his firm a profit or loss of a cent per trade is indeed typical. Prediction, however, is probabilistic, and so many trades lead to losses. This interviewee reported that his firm’s trades were profitable only around 53 percent of the time, which implies an average profit of around 0.06 cents per share traded. When I prompted another interviewee with a higher estimate (a “fifth of a cent per share”), he corrected me:

> Oh, I wish it was that big! There’s not that much, it’s even, yeah, I mean five mils [0.05 cents per share traded], ten mils [0.1 cents], that sort of thing. (interviewee AH)[23]

There was broad agreement amongst those interviewees who were prepared to discuss HFT’s profits — and amongst those in “dark” venues who monitor those profits — that a tenth of a cent per share traded was a roughly accurate indicator of

[23] “Mil” is the term employed in U.S. share trading for a hundredth of a cent.

the order of magnitude of those profits.[24] However, the more recent of the interviews (e.g., BH) suggested that average profit has slipped to around a twentieth of a cent.

Given the controversy surrounding HFT, and in particular the widespread accusation that it preys upon execution algorithms, it might be that interviewees were deliberately underestimating HFT's profitability. In addition, there may be response bias: perhaps the HFT firms that could not be identified or at which no-one could be persuaded to be interviewed were disproportionately profitable. Higher figures for profit rates can indeed be found in the financial-economics literature, in a calculation by one of Lewis's (2014) interviewees, and in the published accounts of Knight Capital, the one major U.S. HFT firm that (prior to its recent takeover by HFT market-maker Getco, following the disaster that struck Knight's automated trading) reported publicly. However, for contingent reasons those figures may be unrepresentatively high.[25] A press report (Massoudi and Mackenzie 2013a) is consistent with recent interviewees' suggestions of current profits of around a twentieth of a cent per share, not a tenth.

[24] E.g. interviewee AK: "… you're operating on making a fraction of a penny per trade … tenth of a penny per share."

[25] The HFT profit figures in Brogaard, Hendershott, and Riordan (2013) equate to an average trading revenue net of fees of around 0.4 cents per share traded. However, they have no data on costs other than fees and their data (described in note 7 above) are mostly for 2008-9, and interviewees reported that period, especially 2008, to be years of exceptionally high HFT profits. An experiment to "neutralize" HFT described in Lewis (2014, pp. 50-52) was estimated by those conducting it to have saved the firm in question $29,000 on a ten-million share order, or 0.29 cents per share traded. Again, though, that experiment seems to have been conducted in 2009, and provides an estimate only of potential HFT revenues, not profits net of costs. Knight's annual reports for 2009-11 (e.g., Knight Capital Group, Inc. 2012) contain data on market-making revenues and expenses that suggest profitability in the region of 0.14-0.19 cents per share traded. However, Knight's activities were broader than HFT (e.g., execution of retail orders), which may explain these relatively high profit rates.

Certainly, my fieldwork impressions were not of great prosperity. Seven HFT interviewees were interviewed twice; by the time (around a year later) the second interview took place, two had lost their jobs. When interviewing at HFT firms, I was sometimes taken on a brief tour of their offices, and often it was possible to see these when arriving for and leaving the interview (most HFT firms' premises are not large). Especially in the later interviews, it was quite common to see several — sometimes many — unoccupied desks. For instance, I visited one large HFT trading room in both March 2012 and May 2013. The number of occupants had visibly shrunk between the two visits, and that impression was confirmed by two interviewees. However, HFTs still transacted enormous quantities of shares. Even in 2013, with overall U.S. share-trading volumes having shrunk markedly from their 2008-9 peaks, and — if the above analysis is correct — with HFTs in aggregate generating only modest profits, HFTs were still buying or selling around 5 billion shares a day, and were in all likelihood a party to the majority of transactions.[26]

THE HINGE: HFT, TRADING VENUES, AND REGULATION

The rise of HFT in U.S. share trading to its current central role came about from the interaction of three ecologies: HFT itself; trading venues (especially new venues called ECNs, or electronic communications networks, a series of which were created

---

[26] The average daily volume of U.S. shares traded in 2013 was around 5 billion (Angel, Harris, and Spatt 2013, p. 4), but each transaction involves both a buyer and a seller, so each trading day around 5 billion shares were bought and 5 billion sold. A reasonable estimate of the HFT share of those purchases and sales is 50-55 percent (e.g., Table 1 and Mackenzie 2012), meaning that the daily total of purchases and sales by HFTs was around 5 billion. If HFT profitability was on average 0.1 cents per share bought or sold, this means that HFTs trading U.S. shares made in aggregate around $5 million per trading day, or $1.25 billion per year. This figure is not large when compared, e.g., with the profits of even a single large bank.

from the mid-1990s onwards); and regulation. This rise, however, has taken place in a context deeply shaped by a decision in the late 1970s about how trading on spatially separate venues should be linked.

Let us begin with regulation. It is indeed an ecology in Abbott's sense. The U.S. financial markets had and have not one regulator but several, including currently seven Federal bodies (the Federal Reserve, Federal Deposit Insurance Corporation, Office of the Comptroller of the Currency, Office of Thrift Supervision, Consumer Financial Protection Bureau, Commodity Futures Trading Commission [CFTC], and Securities and Exchange Commission) and also state regulators such as New York's Department of Financial Services.[27] In consequence, there have historically been both overlaps and gaps in regulatory jurisdiction, and episodes of "turf warfare" (e.g., between the SEC and CFTC: see author ref.). While among Federal regulators share trading has been the largely uncontested terrain of the SEC, the latter shared broader jurisdiction over it with designated "self-regulatory organizations" (notably the New York Stock Exchange and NASD, the National Association of Securities Dealers, which ran NASDAQ).[28] Furthermore, with share trading prominent in American culture, Congress and the executive branch paid far more attention over the decades to shares than to the trading of other financial instruments. That attention has also been a force buffeting the SEC, whose Chair and Commissioners are appointed by the President, subject to Senate ratification, and usually have explicit party-political affiliations. It is noteworthy, for example, that

[27] For example, the government official most vocal in criticism of HFT is not a Federal official, but New York's Attorney General, Eric Schneiderman.

[28] In 2007, NYSE and NASD regulatory functions were merged into the Financial Industry Regulatory Authority (FINRA).

while there have been a plethora of SEC efforts to reform share trading, those efforts are far sparser in the SEC's other main regulatory domain, bonds.

The SEC was a quintessential New Deal institution. It was created in 1934, in the face of the Great Depression, under a President who had declared in his inaugural address that "the money chargers have fled from their high seats in the temple of our civilization" (Seligman 1982, p. 29). The SEC's establishment followed the searing political theater of the Senate Banking Committee hearings, which had exposed pervasive Wall Street wrongdoing. (Amongst dramatic moments was the unscheduled testimony on April 26, 1932 of Fiorello LaGuardia, soon to be Mayor of New York, accompanied by a phalanx of New York police officers, two of whom carried a trunk full of documents, many of them cancelled checks: the evidence of bribes paid by intermediary turned whistle-blower A. Newton Plummer to reporters to write false stories concerning stocks in which his patrons had an interest.)[29]

A persistent theme in the SEC's efforts to reform share trading was its suspicion of the two main self-regulatory organizations with which it shared jurisdiction. The New York Stock Exchange (NYSE) and NASDAQ were in effect a duopoly: companies could choose on which of the two to list their shares, but thereafter those shares traded almost exclusively on the chosen venue. Regional stock exchanges, for example in Boston, Philadelphia, and San Francisco, could trade NYSE shares, but were generally not fully effective competitors to the NYSE. A NYSE "specialist" (market-maker) enjoyed a near-monopoly position in the trading of the stock for which he was responsible (both specialists and floor brokers were nearly always men), and while NASDAQ broker-dealers ostensibly competed with

[29] Seligman (1982, p. 16). LaGuardia's testimony is available at http://archive.org/stream/stockexchangepra02unit/stockexchangepra02unit_djvu.txt, accessed December 27, 2013.

each other there was sometimes tacit collusion among them (see author ref.). These SEC concerns resonated within the political system. In the Securities Acts Amendments of 1975, Congress altered the legislation that had created the SEC, with the goals "to remove barriers to competition" and "to remove impediments to and perfect the mechanisms of a national market system for securities."[30]

How, though, should that "national market system" be designed? Put more broadly, how should economic activity be coordinated across space? Although its consequences could not have been predicted when that decision was taken in the late 1970s, the solution that was chosen still shapes U.S. share trading. One solution would have been to gather economic activity together in a single place, a proposal that became known as the "hard CLOB," or Composite Limit-Order Book, a central, national, electronic order book into which all brokers, dealers or trading-floor market-makers, "wherever physically located," would enter their bids and offers "on an equal competitive footing" (Seligman 1982, p. 521). However, although the Cincinnati Stock Exchange successfully experimented in the late 1970s with an electronic order book, the other exchanges saw the "hard CLOB" as a threat and successfully promoted a different solution, a design for a national market system in which economic activities would continue to be spatially separated (and exchanges would continue to function much as before) but would be linked by a technological network: the Intermarket Trading System, launched in 1978.

The Intermarket Trading System linked NYSE, the American Exchange, and the regional exchanges (NASDAQ was not part of it until 2000). It operated in conjunction with the Consolidated Quotation System, also launched in 1978, which

[30] The Securities Acts Amendments are available at http://www.gpo.gov/fdsys/pkg/STATUTE-89/pdf/STATUTE-89-Pg97.pdf, accessed January 19, 2014.

disseminated information on bids and offers available on the different exchanges. If a broker or “specialist” on an exchange trading floor could see a superior price quotation available on another exchange, that quotation was “protected.” He was not supposed to “trade through” it by dealing on his own exchange at an inferior price, but had to use the Intermarket Trading System to send a “commitment to trade” to the relevant specialist on the exchange with the better price. That specialist then had a set time period — as late as 2002, 30 seconds — to decide whether or not to trade (Hendershott and Jones 2005). If prices were moving fast, that gave the specialist valuable time to see if they would rise or fall (the provision was in effect a version of what in foreign exchange is called “last look”: see below). The Intermarket Trading System thus never threatened the dominance of the New York Stock Exchange’s specialists. As late as 2005, 80 percent of trading of NYSE-listed shares was on NYSE: see Angel, Harris, and Spatt (2013, p. 20).

The episode indicated the limits on SEC’s capacity to impose its will on the self-regulatory organizations and entrenched trading venues. More profound change required a shift *within* the ecology of trading venues. Central to that shift was Island, a new trading venue set up in 1995, originally in the low-status fringes of the U.S. financial markets. Island at first catered primarily for traders known to their established rivals as “SOES bandits.” (SOES was NASDAQ’s automated Small Order Execution System, and “bandits” used it, for example, to pick off NASDAQ broker-dealers’ stale price quotes.) Island made it possible for “bandits” and other “day traders” also to trade directly with each other. Its fees were very low, and its order book was visible to anyone trading on the system (unlike, e.g., the NYSE order book, to which specialists had privileged access). Island’s matching engine was simple and ultrafast, and news of all changes in the order book was disseminated by a specially designed fast datafeed called ITCH; another specialized computer protocol, OUCH, facilitated rapid submission of orders and cancellations of orders. While on NASDAQ, NYSE, and other U.S. venues the "tick size" — the minimum increment of price — was an eighth or a sixteenth of a dollar, Island's tick size was 1/256th of a dollar, making it possible for market-makers on Island to undercut their established counterparts by small (but, from a SOES bandit's or other day trader's viewpoint, economically important) amounts. Market-making was also encouraged by the "rebates" described above; Island was the first venue to introduce rebates.[31]

These features of Island reflected its immediate context. Its developers worked in 50 Broad Street in lower Manhattan, a building also occupied by two of the "bandit" firms for which it catered; one of those firms, Datek, was its original financial backer. The top priority of SOES bandits and many other day traders was speed[32] — hence Island's emphasis on that — and Island's developers (especially its chief architect, Josh Levine) had a libertarian bent, a distaste for oligopolies such as that of NASDAQ's broker-dealers, and a strong commitment to "democratizing" markets: hence the low fees, publicly visible order book, and the small tick size that made it possible to undercut broker-dealers.

Although originally a reflection of these local priorities, those features of Island — which were also adopted at least to a degree by the later ECNs (electronic communications networks) that had to compete with it — came to act as a "hinge" in Abbott's sense, linking developments in trading venues to HFT on the one side, and to regulation on the other. The linkage to HFT was straightforward. It *was* possible

[31] The development of Island is treated in more detail in (author ref.), so that development is simply summarized here.
[32] A survey of such traders by Bernstein & Co. in 2000 "found that 58 percent … rate immediacy of execution as more important than a favorable price" (Blume 2000, p. 9).

prior to Island to conduct algorithmic trading analogous to today's HFT, but it was difficult. Thus HFT interviewee BE recalls what happened to algorithm-generated orders submitted to the New York Stock Exchange via its supposedly automated "SuperDot" order entry system. They were routed automatically to the appropriate specialist's booth on the NYSE floor, but the execution of these orders was controlled manually by the specialist. Even at the best of times, the execution of an order or the cancellation of an order took several seconds, and sometimes the execution of an order would be delayed for much longer, even if there seemed to be matching orders already in the book. "Thirty seconds would go by; sixty seconds would go by." This interviewee came to detect a pattern in such delays and inferred a cause: "somebody's coming to the market with a big buy order. The specialist knows that the stock is going to run up and basically he would freeze his book …" If you were "very agile," some of these frustrations could be turned into opportunities,[33] but "you had to basically put up with those kind of things … you had to learn to live within the realities that you confronted."

On Island, those "realities" were utterly different: a marketable order received via OUCH by its matching engine would be executed in around two milliseconds (interviewee AG1). Although this speed, and Island's other features, were not originally designed to facilitate HFT (as noted, Island's original clientele were manual day traders), they had that effect. Fast matching also motivated what later became HFT's characteristic spatial feature, the co-location of the servers hosting trading algorithms in the same building as the server running the matching engine.

---

[33] If execution of a buy order was delayed in this way, "[y]ou could all but predict" it was because a large buy order was being executed on NYSE. So interviewee BE's firm would sometimes quickly place a bid for the same shares on Instinet (an early electronic venue designed for use by institutional investors) to benefit from the coming price rise.

Transmission delays of a few milliseconds in fiber-optic cables were not salient when matching took several seconds, but with two-millisecond matching a HFT algorithm would be badly disadvantaged unless it was running on a server next to Island's matching engine in the basement of 50 Broad Street.

The "hinge" provided by Island's (and the other ECNs') features also connected developments in trading venues to regulation. Those features made possible what Congress and the SEC had declared they wanted, but had largely failed to bring about: effective competition among trading venues. Using Island and the other ECNs, electronic market-makers could routinely undercut their traditional counterparts, and their doing so made Island and the other ECNs in many respects more attractive places to trade than the traditional venues. Under Arthur Levitt, appointed SEC Chair by Bill Clinton in 1993, the SEC went some way to seizing the opportunity presented by this change in the ecology of trading venues. It helped the ECNs gather momentum, for example with the SEC's new Order Handling Rules, introduced in 1997, which forced NASDAQ's broker-dealers to display ECN prices to their customers when these prices were better than the broker-dealers' own quotes. Amongst the features of Island that the SEC helped generalize was tick sizes much smaller than the traditional eighths of a dollar: in 2000, the SEC imposed "decimalization" (the pricing of shares in dollars and cents). As on Island — but this time across all share trading in the U.S. — a small tick size helped HFT market-makers undercut their established rivals.

The linkage between HFT, Island and the other ECNs, and the SEC was indeed a hinge, not an alliance. Though the SEC's reforms helped facilitate HFT, I know of no evidence that the SEC intended to promote it, or even that SEC officials were aware of its existence (prior to 2005, HFT received almost no publicity even in the specialist financial press, and it was 2009 before its existence became widely known). Nor was there any meeting of minds between the SEC and Island, which was fiercely libertarian and deeply sceptical of the virtues of externally-imposed regulation (see author ref.). Furthermore, the SEC's response to the innovation facilitated by the hinge was a path-dependent response, constrained by the legacy of its 1970s' compromise with the exchanges, the slow Intermarket Trading System. "[U]nder pressure from the exchanges" (interviewee AQ), the SEC insisted in 2002 that Island must join that System. Island "couldn't operate in that world" (AQ), and "went dark," avoiding the imposition by making its order book invisible, in effect becoming a dark pool (Hendershott and Jones 2005). The episode was the first clear sign of the tension between the fast trading unleashed by the "hinge" and the legacy of the 1970s' compromise. It was a harbinger of a contradiction that persists.

REGULATION NMS AND INTERMARKET SWEEP ORDERS

The 1970s' decision to develop the Intermarket Trading System, not the centralized hard CLOB, and the emergence of the "hinge" connecting HFT, trading venues and regulation continue to shape today's share trading in the U.S., but in a way the contradictions of which have become ever more evident. Thirty years of conflict over how to structure share trading in the U.S. culminated in a 2005 measure that still governs that trading: Regulation NMS [National Market System] (Securities and Exchange Commission 2005). On the face of it, "Reg NMS" (as it is universally known) — along with the SEC's earlier measures such as decimalization —

represented an unalloyed triumph of the processes set in motion by the hinge. Reg NMS swept aside the most obvious barrier that the slow Intermarket Trading System placed in the way of fast, electronic trading: under Reg NMS, if a quotation was available only from a human being on a trading floor, it was no longer "protected" but could now freely be "traded through." Only bids or offers that could be hit or lifted electronically and near-instantaneously were protected. The SEC's goal of effective competition to NYSE and its specialists was finally achieved: in four short years from 2005, NYSE's share of the trading of NYSE-listed stocks fell from 80 percent to just over 20 percent (Angel, Harris, and Spatt 2013, p. 20).

Simultaneously, however, Reg NMS still echoed the compromise between the SEC and the exchanges that gave birth to the Intermarket Trading System: in its structure, in the way in which it coordinates economic activity across space, Reg NMS is closer to that system than to the defeated alternative, the single, national Consolidated Limit Order Book. (Reviving the latter proposal would have been very hard politically: even the new venues such as Island opposed it.) Under Reg NMS, trading venues finally did compete with each other — indeed, they do so fiercely — but the form of that competition is not integration into a single order book. Rather, each venue still has its own order book, and how it operates that order book is governed, as with the Intermarket Trading System (only more rigidly so), by rules protecting quotations and prohibiting trade throughs. Consider, for example, the set of order books shown in Figure 1, in which what in Reg NMS is called the "national best offer" of Astoria Financial shares is $7.75. Offers at that price are "protected." Reg NMS prohibits any venue from selling Astoria shares at any price higher than $7.75: to do so would be a prohibited "trade through." Reg NMS similarly prohibits "locking" another venue or venues. Suppose NASDAQ, for instance, received an order to buy 1,000 Astoria shares at $7.75. It can execute 488 shares against the offers in its order book, but is prohibited from posting in its order book a bid to buy the remaining 512 at $7.75. The rationale is that "protected" $7.75 offers are still available on other venues, so the bid can be executed on those venues and must be forwarded to them. To post the bid on NASDAQ would "lock" those other venues, in the terminology of Reg NMS.

These Reg NMS "order protection" rules — largely inherited from the Intermarket Trading System — force the designers of a matching engine to add (either in the engine itself or in separate software) an algorithm that checks whether a new order *can* be executed or entered into the order book, or whether it violates the prohibitions on "trading through" or "locking" protected quotations. This is done by checking the characteristics of the order against the national best bid and offer as determined by the multi-venue datafeeds known as the "consolidated tape."[34]

The checking algorithm has major effects on how HFT algorithms can operate. First, checking takes time, slowing order entry and execution. Second, it is often checking against a past state of the world: the consolidated tape is slower than venues' direct datafeeds transmitted via microwave links, lasers, or the fastest fiber-optic cables, and so an HFT algorithm may "know" that the "protected" bids or offers on the tape no longer exist: they have been executed against or cancelled. (Reg NMS, we might say, is implicitly Newtonian — it implicitly presupposes that instantaneous communication across space is possible — while HFT operates in an

[34] The national best bid and offer for each stock are calculated continuously by the U.S.'s two Securities Information Processor computer systems (one located in Mahwah, NJ; the other in Carteret, NJ). Each SIP, as they are known, receives quote and transaction data from all the exchanges trading the stocks for which it is responsible, and calculates and disseminates the national best bid and offer. Although the SIPs have been improved greatly in recent years, the process of transmission, processing, and retransmission makes these datafeeds lag behind raw datafeeds direct from trading venues' matching engines.

Einsteinian world in which the time that communication takes, even at the speed of light, is salient.) Third, the checking of compliance with Reg NMS constrains "aggressive" HFT algorithms that seek to "sweep the book" (hit multiple existing bids or lift multiple existing offers at multiple prices), because such purchases or sales will be delayed until they no longer appear to trade through the national best bid or offer. Fourth, Reg NMS also constrains the activities of liquidity-making HFT algorithms, because the entry of their orders into a venue's order book will often be delayed until they no longer "lock" other venues. (Delay is a problem for a liquidity-making algorithm, because for such an algorithm to make a profit, at least some of its bids or offers must be executed against, but the chances of this happening depend on where they are in the queue of other bids or offers at the same price. In share-trading matching engines these queues operate on a time-priority basis: the first order to be entered into the book is the first to be executed.)

Reg NMS, however, makes provision for a special category of order that a venue's matching engine can execute or enter directly into its order book without invoking the checking routine: an Intermarket Sweep Order or ISO.[35] An ISO is an order bearing a computerized flag indicating that the firm submitting it has also sent orders that will execute against and thus remove from the order books of all other trading venues any protected quotations that would otherwise be traded through or locked by the order bearing the flag. The SEC seems to have built provision for ISOs into Reg NMS to avoid the latter delaying large orders from big investors (Securities and Exchange Commission 2005, p. 37523), but these orders have become crucial to the successful practice of HFT: as well as avoiding delays caused by Reg NMS's

---

[35] The Intermarket Sweep Order exception to Reg NMS is defined in section 242.600(b)(30) of Securities and Exchange Commission (2005).

"Newtonian" rules, use of the ISO flag can speed order entry or execution simply because the trading venue's computer system need not invoke the routine that checks an order's consistency with Reg NMS.

Intermarket Sweep Orders are the most important single way in which the contradiction between the processes unleashed by the hinge and continuing path-dependence — the legacy of the Intermarket Trading System — are resolved in HFT practice. In effect, ISOs allow Einsteinian actors (HFT algorithms with access to nearly speed-of-light market datafeeds) to circumvent the effects of Reg NMS's Newtonian regulation. All the HFT interviewees who were prepared to talk about ISOs said they were important. Interviewee BE, for example, reported that they "created speed opportunities": if you didn't use them, "you'd be behind the queue." "[A] large amount of wealth transfer happens here," said interviewee AF. However, not every algorithm (until recently,[36] not even every HFT algorithm) can employ an ISO flag: only registered broker-dealers are permitted to use it (and that registration brings heavy extra costs via compliance requirements), although broker-dealers can delegate the right to use it to trusted customers such as experienced HFTs. Those whose algorithms cannot use the flag are disadvantaged, said AF (whose HFT firm is not a broker-dealer): a HFT "can send an order with a ISO flag to post first, then later all those investors who were trying to post will post behind you with inferior time priority." Drawing on data from 2010, Madhavan (2012) reports that 28 percent of U.S. share trading, and 21 percent of trading in exchange-traded funds, involved the use of ISOs; McInish, Upson, and Wood (2012) report higher incidences (40-45 percent) of ISO use in the trading of stocks in the S&P 500.

[36] Interviewee AP, in October 2012, reported that "the … problem with people [HFTs] not being able to send ISOs has been solved."

The use of Intermarket Sweep Order flags, however, is only the most evident manifestation of the contradiction between the hinge and path-dependent regulation. The behavior of HFT (and other) algorithms trading U.S. shares is deeply shaped by the detail of Reg NMS. Much — though not all — of this shaping is via what a liquidity-making algorithm with access to fast, direct datafeeds and predictive capacity needs to do, in order to ensure the maximum benefit from that access and this capacity by gaining the most favorable possible position in the time-priority queue. This issue erupted into controversy in 2012, when algorithmic trader Haim Bodek, founder of Stamford, Connecticut options trading firm Trading Machines LLC, told the SEC and *Wall Street Journal* that trading venues were making specialized types of order available to help liquidity-making HFTs optimize their positions in time-priority queues subject to Reg NMS (Patterson and Strasburg 2012).[37] HFT interviewees' responses to Bodek's critique were nuanced. On the one hand, they denied that these order types were secret. "If you read the [*Wall Street*] *Journal* you were misled," said interviewee AQ: a "fairly clear description of how they [the specialized order types in question] operate" was available. "All those things are out

[37] Option-trading algorithms such as Bodek's hedge their positions by buying or selling shares, and traditionally have done that by taking liquidity. Trading Machines's strategy was to do it by making liquidity (Patterson 2012, p. 23), which, if could be done in a timely way, would be economically advantageous for the reasons discussed above. However, the firm repeatedly found that its algorithms' orders were not being executed, presumably because of unfavorable queue positions. Much of the controversy focused on the "Hide Not Slide" order on the trading venue Direct Edge. "Sliding" refers to the procedure of repricing orders so that they can be displayed in a way that complies with Reg NMS (e.g., reducing the price on a bid so that it no longer locks another market). An order that has been "slid" is then repriced at the original price if that market moves in such a way that the order's display becomes consistent with Reg NMS, but normally its position in the time-priority queue is determined by when this repricing takes place. In contrast, the time priority for a Hide Not Slide order is determined by the time of the order's initial placement; it is entered into the order book but not displayed until its display is permissible under Reg NMS, hence the name Hide Not Slide. For an explanation — published at the time of Direct Edge's introduction of Hide Not Slide — see Anon (2009).

there," said BF. On the other, however, no HFT interviewee prepared to speak about the issue denied that specialized order types were important. "They're hugely important," said interviewee AP: "it's been around forever, they [journalists] just found out about it." "It does go on," said AR.

DARK POOLS AND ALGORITHMIC BOUNDARY WORK

"Technical" solutions such as the use of Intermarket Sweep Orders and specialized order types are one way tensions and contradictions are resolved in practice. Another is boundary work. As noted above, common accusations are that HFT algorithms prey upon institutional investors' execution algorithms — "You're basically getting your face ripped off," said an interviewee whose firm supplies these algorithms — and that stock trading is "rigged" in favor of predatory algorithmic behaviour (see, e.g., Lewis 2014). Indeed, the underlying issue of contested legitimacy resonates within HFT itself. Those HFT firms whose algorithms specialize in liquidity-making can, as also noted above, believe that those algorithms act more "morally" than algorithms that take liquidity, and claim for themselves the legitimacy of the traditional role of the market-maker. By taking on the identity of "electronic market-maker," HFT firms thus sometimes seek to draw within the domain of economic action a boundary of the kind described by Zelizer (2012). For example, when I mentioned the name of another HFT firm to the head of communications at interviewee AC's firm, she distanced that firm from her's: her firm's business was "more pure-play market-making." Interviewee AC agreed, distinguishing his firm from others that "trade in any style that looks like it might make some money." Interviewee

AQ, from a different firm, drew the same distinction: “We’re an electronic market-maker. We unfortunately fit under the definition of high-frequency trading.”

Drawing a boundary between electronic market-making and other forms of HFT would be easier if market-making algorithms only made liquidity, and never took liquidity. Sometimes, however, a market-making algorithm cannot fully control risk by “shading” its bids and offers, and needs to take liquidity to reduce its risk. No market-making HFT firm of which I am aware bans its algorithms from ever taking liquidity, and junior traders do not always feel the boundary impinging on their design and use of algorithms: “you can really do anything,” said interviewee AW1, who worked for a firm that positioned itself as a market-maker.

Another form of algorithmic action that is a candidate for being “bar[red] … as inappropriate” (Zelizer 2012, p. 145) is algo-sniffing: a HFT algorithm setting out to detect and exploit execution algorithms. Indeed, algo-sniffing is sometimes explicitly disavowed by HFTs that position themselves as market-makers: it’s “not something we’ve done,” said interviewee AC; “we dismissed the idea.” “We choose not to do it, but someone like us could do it,” said interviewee AQ. However, AP’s warning, quoted above, about the difficulty of humans being certain what their algorithms are really doing indicates the difficulty of barring algo-sniffing. All the HFT firms in which interviewees were prepared to go into this level of detail use the dynamics of order books as a source of prediction, and some firms, rather than just programming their algorithms to detect patterns identified by human beings, also employ machine learning techniques, in which the algorithm itself searches for patterns with predictive power. It is not clear that even the algorithm’s owner can then be certain that its success is not actually based on algo-sniffing, rather than, for example, detecting a less specific form of price “momentum.” (Interviewee AQ, who was strongly committed to his firm's identity as an electronic market-maker, said the firm went as far as to eschew machine learning: "[w]e have no pattern recognition [in the firm's algorithms].")

Furthermore, the validity of seeking to draw a boundary between electronic market-making and "aggressive," "opportunistic," or "algo-sniffing" HFT is fiercely contested. Interviewee BI, for example, saw liquidity-taking (and not just for risk management purposes) as entirely consistent with a market-making role, preferring to refer to the activity as "liquidity-satisfying." Many U.S. HFT interviewees were in varying degrees libertarian, viewing algo-sniffing as just as legitimate as other forms of behavior within the law. ("[T]he guy who's trying to hide the supply-demand imbalance [by using an execution algorithm]," said a broker: "why is he any better of a human being than the person trying to discover [that imbalance]" by running an "algo-sniffing strategy"?) Some particularly strong libertarian interviewees even defended "spoofing" (adding spurious orders to order books so as to deceive algorithms that use order-book imbalances as a basis of prediction).

Despite all the difficulties that surround them, attempts within HFT to distinguish between market-making and other forms of HFT resonate with efforts by the owners of trading venues to draw a similar boundary. In 2011, "Light Pool," a new "lit" venue — a venue with an order book visible to participants — was created by Credit Suisse, with this boundary work as its chief rationale. However, the main context of this boundary work is "dark pools," which are trading venues whose order books are not visible to those trading on them. The two earliest dark pools were Posit (set up in 1987) and Liquidnet (set up in 1999), and they were followed by a series of dark pools established by major investment banks. The first was Credit

Suisse's Crossfinder, launched in 2006; others such as Goldman Sachs's Sigma X, Lehman Brothers' (now Barclays') LX, and UBS's ATS soon followed.

The goal of most dark pools is to be a venue in which a "natural" seeking to buy a big block of shares can trade with another "natural" seeking to sell a corresponding block, without the existence of either order being visible to professional traders such as HFTs. (A "natural" is the industry term for an institutional investor genuinely wishing to buy or sell in large quantity; the term can of course carry the connotation that other motivations for buying or selling are unnatural.) However, there may often simply not be a "natural" wanting to buy when another wants to sell (or vice versa), so it can be difficult to achieve adequate liquidity in a dark pool without allowing professional traders to participate. By 2013, around 15 percent of U.S. share trading was in dark pools (Angel, Harris, and Spatt 2013, p. 22), and by then most of them no longer catered simply for large trades between "naturals." HFTs had joined them, and the average size of trades in most dark pools — at around 200 shares — was no larger than those in lit markets (Massoudi and Mackenzie 2013b).

The defining characteristic of a dark pool — the invisible order book — impinges on those HFT practices of prediction that depend on the dynamics of that specific order book, but other sources of prediction remain available and HFT market-making remains entirely feasible; indeed, it may well be necessary to adequate dark-pool liquidity. However, the wider controversy over HFT interacts with frequent suspicions that information "leaks" from dark pools, making some of them "toxic," as interviewee AE put it. Asked what he meant by calling a dark pool "toxic," he replied: "I mean that there's high-frequency trading dudes in there." There are fears that if HFT algorithms can detect large orders in dark pools — for instance, by "pinging" the order book (repeatedly sending in tiny orders to discover whether they are executed) — they can then trade profitably in lit markets. For example, a typical way of executing a large institutional order is to make as many purchases or sales as possible in dark pools, and then execute the remainder in lit markets. A particular fear, therefore, is that an algorithm that can detect the order in a dark pool, at least probabilistically, can position itself to profit when the purchase or sales in lit markets begin.

If a dark pool is seen as "toxic," institutional investors will not want to use it, so it is important to dark-pool operators to convince those investors that their pools control algorithmic behavior that leads to toxicity. As well as operators giving markets technical features designed to make common HFT techniques more difficult,[38] they also directly monitor the behavior of the algorithms and other participants trading in them. Seven interviewees involved with dark pools described measures they take, and there are also useful accounts in the specialist trade press (Mehta 2011; Chapman 2012a&b). One venue is confident enough of its liquidity that it does not allow known HFTs to join it. Another classifies participants into three categories ("contributors," "neutral," and "opportunistic") and expels those classed as "opportunistic." Yet another uses a five-point scale of this kind, but does not expel the "opportunistic." Instead, it makes the categorization available to its matching engine, and allows users of the dark pool to choose to restrict the categories of other participants eligible to trade with their algorithms, so that they can avoid the

[38] For example, Light Pool, as a "lit" market, had to accept Intermarket Sweep Orders, but an algorithm seeking to sweep Light Pool had to send its order via the National Stock Exchange, a small trading venue based in Jersey City. "It takes almost half a second," said an interviewee: "that's almost eternity for a high-frequency trader." See Lewis (2014) for a detailed account of how IEX, another new venue that seeks to constrain HFT, attempts to do so.

"opportunistic" if they want. Another venue, owned by a major supplier of execution algorithms to institutional investors, builds the results of its monitoring into those algorithms when these operate in its own dark pool (for example, to help those algorithms avoid other participants that the monitoring suggests will be successful opportunists in particular circumstances).

Given the quantities of data that need analyzed to determine whether a firm's algorithms are "opportunistic," this boundary work is itself largely algorithmic. A common technique — three interviewees reported use of it — is to estimate a firm's short-term profitability, for example by monitoring how the price of the shares it trades move over the second after the consummation of each trade involving it. Too high a profit rate is taken as an indicator of opportunism. Another variable that is used is the proportion of a firm's algorithms' trades that are liquidity-taking; if that figure is high, that too is taken as suggesting opportunism. Interestingly, only two of the seven interviewees whose trading venues engaged in this boundary work viewed it explicitly as moral in nature. The view that "this is business," as one interviewee put it, was more common. (Other motivations for surveillance and for boundary work that were cited were simply that "clients do want it"; that it was important to be able to demonstrate that when an investment bank's own trading desks traded in its dark pool — which is not uncommon — their trading was benign; and that in a situation in which dark pools were being heavily criticized by the leaders of "lit" markets such as the New York Stock Exchange, it was vital to be able to show regulators that behavior in dark pools is under "full control.")

As noted above, the boundary that is drawn within HFT as a result of surveillance by dark pool operators seems often to coincide with that drawn by the HFT firms that position themselves as electronic market-makers. As one venue operator put it in interview, in distinguishing algorithmically between "good flow and … bad flow … between the good guys and the bad guys," the "good guys" were "market-makers." The very act of surveillance pushed algorithmic behavior in this direction, noted another. Some HFTs rejected surveillance designed to curb their algorithms: "I had one prominent high-frequency shop that on principle refused [to join the venue managed by this interviewee], saying that … the very concept demonized high-frequency trading and therefore we could f-off." Other HFTs, however, embraced surveillance and expressed willingness to modify the behavior of their algorithms to meet its demands. "Tell me how I do," a representative of one electronic market-maker told *Traders Magazine* (Chapman 2012a), "and I'll adjust. … I want to be scored. Everyone should be scored."

LAST LOOK AND THE LINKED ECOLOGIES OF FOREIGN EXCHANGE

Foreign exchange is another domain in which there are restrictions on and surveillance of HFTs, and their consequences for the practices of HFT are greater than those of the analogous measures in share-trading dark pools. Traditionally, currencies were not traded on an exchange, but "over-the-counter," in other words directly between institutions.[39] As on NASDAQ prior to the incursion of the ECNs, foreign exchange was — and to a significant extent remains — a "dealer market." Most market participants — more minor banks, other corporations, hedge funds, and institutional investors — did not and largely do not trade among themselves: they traded only with dealers, by accepting the latter's quotes (in foreign exchange, the

[39] Currency futures, however, were and are exchange-traded, mainly on the Chicago Mercantile Exchange.

main dealers are all big banks). Dealers, however, did and do trade with each other, via interdealer brokers, messaging systems (see Knorr Cetina and Bruegger 2002), or one or other of the two main interdealer electronic-trading venues, Reuters and EBS (Electronic Broker System). The latter was created in the early 1990s by a consortium of dealer banks, and bought in 2006 by the interdealer broker ICAP (Knorr Cetina 2007 gives an excellent account of trading on EBS).

Just as first NASDAQ and then the New York Stock Exchange were challenged by the ECNs, so this dealer-market configuration in foreign exchange was challenged at the end of the 1990s and in the early 2000s by new electronic trading venues modelled on the ECNs. These new venues had something of the character of what Abbott calls "avatars": they were the result of attempts to create incarnations of institutions from one domain (share trading) in a different domain, foreign exchange. However, "the internal forces of competition in the avatar's ecology tend to drive the avatar in directions unforeseen ahead of time" (Abbott 2005, p. 269). Initially, said an interviewee, "people believed that this [the creation of ECN-like trading venues in foreign exchange] was going to force the banks into the new paradigm," as had happened in share trading. However, the dealer banks "were able to say no and not participate" in developments in trading venues that threatened their interests, and without their support and the large volumes of liquidity they could provide, it was difficult for a new foreign-exchange trading venue to thrive.

So confrontation between the new electronic venues and banks had to give way to collaboration. At one new venue, "the original investors fired the CEO," said an interviewee, and his replacement "changed the business model … he befriended the banks … and work[ed] with them over the course of the years." As automation of foreign-exchange trading took place, therefore, it was often shaped by the interests and preferences of the big banks. For example, equivalents of Island's ultra-fast ITCH datafeed and OUCH order-entry protocol were not adopted widely in foreign exchange. Instead, the protocol mainly used in foreign-exchange trading was and is FIX, which was much slower (the above interviewee described it as "verbose" compared to the "compact, efficient" ITCH and OUCH), but widely used in banking and already familiar to banks' information-technology departments.

In share trading, the three crucial linked ecologies were HFT, trading venues, and regulation. As HFT began in foreign exchange, often introduced by firms that were already trading shares or futures, its practitioners confronted a different configuration of linked ecologies. The ecology of trading venues seemed similar to share trading — two large, established foreign-exchange trading venues (EBS and Reuters); a number of new ECN-like venues; and other venues run by banks — but in foreign exchange the crucial linkages of that ecology were not to regulation but to the big dealer banks. As an interviewee said, "when someone calls or we call someone we want to trade on [his ECN-like venue], we ask them if they have existing bank relationships." In share trading, HFTs can operate relatively independently; indeed, larger HFTs often become broker-dealers in their own right. In foreign exchange, if an HFT is to operate on any scale, it needs a bank to act as its prime broker, facilitating its trading and especially the settlement of its trades.[40]

In foreign exchange (as, indeed, as elsewhere), banks appear ambivalent about HFT, welcoming the income (such as prime brokerage fees) proprietary trading firms bring them, but anxious about HFTs as competitors in trading. Banks' organizational structures do not seem conducive to the creation of pared-down,

[40] HFT firms cannot get full membership, noted interviewee BH, of the international Continuous Linked Settlement system in foreign exchange. They therefore have to work through a bank that is a member of the system.

speed-optimized technical systems. In HFT firms, which are (as noted) usually relatively small, privately-held proprietary trading businesses, trading activities and the development of technological systems are intimately interwoven; often, there is no clear organizational distinction between traders and system developers. In a bank, system developers are not normally managed by those responsible for trading, but are part of a separate IT department. “That’s problematic,” says interviewee BE, who has worked in both HFT firms and a bank. Banks’ IT departments have their own preferred technological styles and priorities. In one bank, said BE, the IT department tried to insist that the HFT system must have a security firewall. “[G]uess what, that firewall takes 50 milliseconds to go through it. I can’t do that: I’m out of the game.” Another interviewee, a technology specialist, talked about the problem in banks of “legacy systems” that “just get ingrained in the very fabric of the organization, … doing it [HFT] in a large investment bank is really hard work.”

The linkages between trading venues in foreign exchange and banks, and the ambivalence of the latter about HFT mean that measures to restrict HFT are more prevalent and harsher in foreign exchange than in share trading. HFTs in foreign exchange that are too profitable — especially those that “pick off” banks’ stale quotes — do not simply face, as in the dark pools discussed in the previous section, admonitory telephone calls or their algorithms being electronically tagged as “opportunistic,” but more brutal sanctions.[41] If they are trading on a bank-owned venue and the bank “figures out what they are doing,” said a foreign-exchange trader, it “normally just cuts them off the day they figure it out: ‘that’s it, your account’s closed here, take your money away and don’t come back.’” Even on an

[41] Expulsions of HFT firms from trading venues do happen in share trading, but are not commonplace: I know of only four that I am confident took place.

ECN-style venue, a HFT whose algorithms are too successful can get frozen out by the larger participants. “ECNs are naturally anonymous,” said interviewee AU, but “most of them if not all provide client ID [a firm’s numerical identifier], and the bank can say: ‘oh, I don’t want to trade with this person because they’re good; let’s turn them off.’” Interviewee BL, for example, reported starting high-frequency trading on a new foreign-exchange venue, “but I got turned off in two days’ time because they said we were too predatory.”

Attempts to create technical obstacles to HFT are also more prominent in foreign exchange. Share trading venues that create obstacles, such as Light Pool, the dark pools also discussed in the last section, or IEX (described by Lewis 2014), are “niche” markets rather than large-scale. In foreign exchange, in contrast, the traditionally biggest venues seek to constrain HFT. Both Reuters and EBS impose a minimum order resting time: on EBS, for example, an order must remain in the electronic order book for a quarter of a second before it can be cancelled. In 2012, EBS reversed an earlier move to reduce minimum price increments, increasing the increment five-fold for some currency pairs, and it also then slowed trading down by replacing its matching algorithm (which was a time-priority algorithm which “queued” orders in essentially the same way as share-trading matching engines do) with an algorithm that collects incoming orders into a batch and then randomizes the order in which it processes them. Speed “is a technology arms race to the bottom, and a huge tax on the industry,” said the new chief executive of EBS. Its preferred customers “are serious players who come to the market to exchange risk — they do not come to race” (Foley 2013). A new bank-backed venue, ParFX, slows down matching in a different way, by delaying each incoming order for a randomized period of between 20 and 80 milliseconds.

There is thus a sharp contrast to the mainstream “lit” trading venues for shares. These compete to be fast, while the different links of the ecology of trading venues for foreign exchange mean that mainstream venues there can believe they can become more attractive by being slow. Most distinctive of all, however, is an institution that has emerged in the algorithmic trading of foreign exchange as a direct result of that sphere’s chief contradiction: that the algorithms of its dominant actors (big banks) tend to be slow. The institution is called “last look.” Before a matching engine consummates a trade involving a participant granted last-look privileges (normally, a major dealer bank), the engine sends a message to the participant’s trading system giving the latter a period of time (“[a]nywhere from five to ten milliseconds, up to a few hundred milliseconds, sometimes up to a few seconds,” reported interviewee AU) in which it can reject the trade. “ECNs don’t usually have a choice” in respect to last look, said AU: “that is where the power of the liquidity provider comes in.” That provider says to the ECN:

> “if I’m not ‘last look,’ I’m not going to provide you liquidity.” So the ECN gets to a certain point where [the] top 15 people in the world [the major dealer banks that] can provide liquidity are asking for last look and if they don’t [grant them last look] they don’t get access to liquidity for [the ECN’s] clients to take. (interviewee AU)

The rationale for last look from a dealer bank’s viewpoint is that it enables the coordination of economic activity across space: the bank will continually be making liquidity (posting bids and offers) on multiple, geographically separate foreign-exchange trading venues. The matching engines of those venues tend to be slower than in share trading, the banks’ systems are usually slower than HFTs’ systems, and (as noted) the dominant communications protocol in foreign exchange, FIX, is slower than its counterparts in shares. So coordination across space is slow. The dealer bank thus faces the risk that all its offers will be lifted simultaneously or its bids hit simultaneously on multiple platforms, or that the prices of those offers and bids will become "stale" and will be picked off before the bank is able to cancel them. A representative of one of the ECN-like venues in foreign exchange thus defends last look as a necessary institution:

> If you're offering the same price at two different venues, you don't really want to deal twice, so one of the main safeguards that last look provides is the ability to validate that the trade is legitimate; that the price is the best you can get at that instant; and, from a risk perspective, that they're not getting taken out on three venues at once and exposing themselves to too much risk. If that happens too many times to the market-maker, they widen out their quotes, which doesn't help anybody. (Smith 2012)

To high-frequency traders who come to foreign exchange from trading shares or futures, "last look" is at best a peculiarity that has to be lived with, at worst a scandal. Interviewee AK took the latter attitude. In trading foreign exchange, he said, "you're a second-class citizen if you're not a bank." His firm had had some success in the high-frequency trading of foreign exchange, but the "last straw" had been when a trade against one of its liquidity-making orders (it was pursuing an electronic market-making strategy) got rejected on "last look" by a participant that had traded against his firm's passive order:

> When we complained about it, they [the ECN] said, "that's our structure: these certain participants get last look on everything." We

> were, like, we're done. You can give us back our $100,000 we have on account … and close our accounts and you'll never see a dime from us again. We're out. (AK2)

Other HFTs took a more accommodating attitude. Interviewee BC, whose firm's background was in futures trading, said that to be successful in HFT in foreign exchange, one had to "develop relationships" with banks, and "to be careful not to be so carnivorous as in the futures." One had to be "nice," and to provide "more friendly [order] flow," rather than simply pursuing maximum profits. Similarly, interviewee AU said of the HFT of foreign exchange:

> It's all about relationships. … You have to build a pretty good relationship to get access to certain things. … FX spot [the trading of foreign exchange for immediate delivery] is an over-the-counter, electronically-traded, relationship business. And often some of the electronic players just focus on the first two components but without the relationship.

CONCLUSION

What becomes of economic sociology when markets and most participants in them are computer algorithms? This paper has put forward a historical, ecological, and "Zelizerian" sociology of algorithmic trading. It has examined high-frequency trading algorithms: the forms of price prediction on which they rely, their interactions with matching engines and execution algorithms, and the liquidity-making or liquidity-taking actions they take. It has described how some high-frequency traders themselves, and some of the venues on which HFTs trade, engage in Zelizerian boundary work, seeking to distinguish appropriate from less appropriate or unacceptable algorithmic economic action, but has also shown how that boundary work is contested. The paper has offered a politically inflected historical sociology of HFT, and has employed a sociotechnical variant of Abbott's (2005) model of "linked ecologies." The paper has shown, for example, how combinations of past events (especially the choice in the late 1970s of a network form of coordination across space in the trading of U.S. shares) and different ecological links (to regulation, in the case of U.S. shares; to the big dealer banks, in the case of foreign exchanges) shape the current practices of algorithmic trading of U.S. shares and foreign exchange, creating radical differences between the two spheres. Coordination of economic activity across space is achieved by different means, and while mainstream trading venues in shares prioritize speed, some in foreign exchange deliberately aim to be slow. The differences between the two spheres both create and are created by different contradictions (in shares, Newtonian regulation of an Einsteinian sphere; in foreign exchange, the dominance by slow actors of a sphere that, technologically, is potentially ultrafast) that lend salience to quite different algorithmic actions: in shares, Intermarket Sweep Orders; in foreign exchange, last look.

I would conjecture that these processes (historical path-dependence, the effects of different patterns of ecological linkages, boundary work) are to be found not just in financial markets but in most — perhaps all — the spheres in which algorithms are economic actors: for example, Poon's (2007, 2009) account of algorithmic credit scoring could be reinterpreted in this way. Nevertheless, the indefinite article in the paper's title is meant seriously. This particular form of the sociology of algorithms is only one of many that are conceivable. For instance, a broader cultural sociology of algorithms is clearly possible. Island, say, was as much a cultural project as a economic one (author ref.). The fears about HFT that have fueled the boundary work discussed in this paper can plausibly be interpreted as expressing two wider cultural anxieties: about out-of-control technology (for which, see Winner 1977) and out-of-control finance (an anxiety that is of course entirely reasonable following the global financial crisis).

To take another example of a different potential form of the sociology of algorithms, the interviewees' comments on the importance of "relationships" to the successful high-frequency trading of foreign exchange, and the importance of the circulation of personnel to the diffusion of HFT techniques, point to the possibility of a network sociology. The form that it needs to take will be affected by the "topological" shift that has taken place,[42] at least in share trading: that shift may mean that network links among technologists and between HFT firms and trading venues are now more important than those among human traders. There were many pointers in the interviews to the importance of network links, in the form of the movement of staff

---

[42] As noted above, HFT in shares began in the 1990s largely enfolded within trading venues that were still places predominantly of human interaction, and those who conducted HFT generally had to accept the features of those venues and that interaction as a fixed environment to which HFT had to adapt: "you had to learn to live within the realities that you confronted" [interviewee BE, as quoted above]. The topological shift is that by 2008, that relationship had reversed: share-trading venues, faced with fierce competition for market share and even survival, had to adapt to HFT, at least as much as vice versa. Venues needed HFT market-makers to provide the tight "spreads" that would attract business, and so had to offer those market-makers co-location, a fast matching engine, a direct datafeed, low fees, rebates, useful order types, etc. In some cases, personnel from HFT firms helped venue staff overcome the resultant technical challenges (see author ref.). In other cases, HFTs provided funding for new trading venues, or even themselves launched them. The new venue BATS was launched in 2005 by the Kansas City HFT firm, Tradebot. In early 2014, BATS merged with another ECN-like exchange, Direct Edge, to form BATS Global Markets, and the new firm is on the brink of becoming the world's largest share-trading venue. Its market share is almost identical to that of NYSE, which has itself been bought by an upstart electronic futures-trading venue, the InterContinental Exchange.

both between HFT firms (this movement, as noted above, is a major mechanism by which HFT techniques diffuse) and between trading venues and HFTs. For example, technical staff from venues can be particularly valuable recruits for HFT firms, bringing with them a detailed understanding of venues' order gateways and matching engines, and also, for example, knowledge of exactly whom at the venue it is best to speak to if the venue is inadvertently catering poorly for the HFT firm in question, for instance if the firm's connection or connections via order gateways to the matching engines seem to be persistently slow.[43]

Ultimately, though, a sociological analysis of algorithmic markets will require not just an extension of existing cultural, network, and other approaches, but new approaches and certainly new methods. How algorithms act will need to be studied directly, rather than — as here — indirectly, via the medium of interviews. That will require data we do not currently have, and public policy measures will almost certainly be needed to ensure those data become available. New ways of analysing these data are needed as well: we need to know how algorithms *inter*act, a topic on which, in respect to the markets covered in this paper, far too little is systematically known.[44]

For now, however, it is worth re-emphasizing this paper's central finding: that there is no teleology to the triumph in U.S. share trading of anonymous, automated, and highly competitive markets, seemingly close to the economists' ideal "perfect

---

[43] Unfortunately, while mining Linkedin profiles would yield data, albeit unsystematic, on network links, the values of the dependent variable of greatest interest (HFT firms' profit rates) are not knowable with any precision.

[44] Much of the controversy over HFT boils down to competing accounts of how HFT algorithms interact (via the matching engines of multiple trading venues) with execution algorithms. The questions that raises are in good part empirical, but the data to answer them are not publicly available, and the likely complexity of the interaction and the sheer volume of data that will eventually need analyzed highlight the need for new methods.

markets”. As already noted, those markets, and the practices of HFT within them, are still shaped deeply by the legacy of a contingent historical event: the decision in the late 1970s about how to coordinate economic activity across space — the decision not to adopt a “Hard CLOB,” but to follow the preference of most of the exchanges and develop an Intermarket Trading System. Furthermore, the comparison between share trading and foreign exchange shows that the market form that has triumphed in mainstream U.S. share trading was not natural or inevitable; instead, it was the result of the particular “hinge” connecting HFT, trading venues, and regulation. Absent that hinge, and with no equivalent of the SEC, foreign exchange trading venues are, as noted, quite different. In addition, as we have also seen, alongside the apparently “perfect” lit markets in shares, and hosting a growing proportion of trading, are “dark” markets whose operators often perform a version of Zelizerian boundary work. These markets contribute, some high-frequency traders complain, to the mainstream lit markets, far from actually being “perfect,” becoming full of what HFT interviewee BH called toxic “exhaust.”[45]

“[M]arkets are not more or less social,” commented Mark Granovetter a decade ago: “They may be more or less personal” (Krippner et al. 2004, p.129). Algorithmic markets in which most actors are themselves algorithms are the most depersonalized of current market forms. Revealing the many ways in which these markets are still social is a crucial challenge for economic sociology. This paper has

[45] As noted above, institutional investors’ execution algorithms typically try first to execute their orders in dark pools, before routing them to lit markets. If a large sell order, say, finds no buyers in dark pools, it probably means that prices are about to fall. The HFT market-making algorithms whose bids are hit when the order reaches the lit markets will thus lose money. “Adverse selection” of this kind seems substantial: at around 0.1 cents per share traded (or perhaps less), HFT profits are lower than rebates (which are, as noted, around 0.3 cents per share traded), indicating that without rebates HFT market-making at current “spreads” between the best bid and best offer would on average lose money.

identified some of those ways, and this conclusion has sketched other possibilities, but many more remain to be discovered. The sociology of algorithms is still in its infancy, but its growth to maturity will surely be a major aspect of economic sociology in the years to come.

REFERENCES

Abbott, Andrew. 2001. *Time Matters: On Theory and Method*. Chicago, University of Chicago Press.

—. 2005. "Linked Ecologies: States and Universities as Environments for Professions." *Sociological Theory* 23:245-274.

—. n.d.. V: Ecologies and Fields, Available at http://home.uchicago.edu/aabbott/Papers/BOURD.pdf, accessed February 1, 2014.

Angel, James J., Lawrence E. Harris, and Chester S. Spatt. 2013. Equity Trading in the 21st Century: An Update, June 21. Available at: http://www.knight.com/newsroom/pdfs/researchCommentary06272013.pdf, accessed January 21, 2014.

Anon. 2009. Direct Edge Launches Hide Not Slide Order, May 28. Available at http://www.thetradenews.com/print.aspx?id=2597, accessed December 31, 2012.

—. 2010. "It's a Screaming Hold." *Financial Times* October 9-10:24.

Arthur, Brian W. 1984. "Competing Technologies and Economic Prediction." *Options* April:10-13.

Baron, Matthew, Jonathan Brogaard, and Andrei Kirilenko. 2012. The Trading Profits of High Frequency Traders, Typescript, November.

Beunza, Daniel, and David Stark. 2004. "Tools of the Trade: The Socio-Technology of Arbitrage in a Wall Street Trading Room." *Industrial and Corporate Change* 13:369-400.

Beunza, Daniel, and Yuval Millo. 2013. Folding: Integrating Algorithms into the Floor of the New York Stock Exchange, Available at http://ssrn.com/abstract=2265849, accessed December 2, 2013.

Blume, Michael E. 2000. The Structure of the US Equity Markets, Available at http://finance.wharton.upenn.edu/~rlwctr/papers/0017.pdf.

Brogaard, Jonathan A., Terrence R. Henderschott, and Ryan Riordan. 2013. High Frequency Trading and Price Discovery, Available at: http://ssrn.com/abstract=1928510, accessed 14 November 2013.

Budish, Eric, Peter Cramton, and John Shim. 2013. The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response, Available at: http://faculty.chicagobooth.edu/eric.budish/research/HFT-FrequentBatchAuctions.pdf, accessed July 8, 2013.

Çalişkan, Koray, and Michel Callon. 2009. "Economization, Part 1: Shifting Attention from the Economy towards Processes of Economization." *Economy and Society* 38:369-398.

—. 2010. "Economization, Part 2: A Research Programme for the Study of Markets." *Economy and Society* 39:1-32.

Chapman, Peter. 2012a. "Barclays Polices Dark Pool," Traders Magazine Online News, April 25. Available at http://www.tradersmagazine.com, accessed April 25, 2012.

Chapman, Peter. 2012b. "Keeping Watch", Traders Magazine Online News, May 7. Available at http://www.tradersmagazine.com, accessed May 7, 2012.

David, Paul A. 1992. "Heroes, Herds and Hysteresis in Technological History: Thomas Edison and 'The Battle of the Systems' Reconsidered." *Industrial and Corporate Change* 1:129-180.

Deleuze, Gilles. 1990. *The Logic of Sense*. New York, Columbia University Press.

Durbin, Michael. 2010. *All About High-Frequency Trading*. New York, McGraw-Hill.

Fligstein, Neil. 1996. "Markets as Politics: A Political-Cultural Approach to Market Institutions." *American Sociological Review* 61:656-673.

—. 2001. *The Architecture of Markets*. Princeton, NJ: Princeton University Press.

Fligstein, Neil, and Doug McAdam. 2012. *A Theory of Fields*. Oxford, Oxford University Press.

Foley, Stephen. 2013. "HFT Platforms Facing 'Speed Limits'." *Financial Times* April 29:22.

Du Gay, Paul, Yuval Millo, and Penelope Tuck. Forthcoming. "Making Government Liquid: Shifts in Governance using Financialisation as a Political Device." *Environment and Planning C.*

Gieryn, Thomas. F. 1999. *Cultural Boundaries of Science: Credibility on the Line*. Chicago, Chicago University Press.

Hasbrouck, Joel. 2003. "Intrading Price Formation in U.S. Equity Index Markets." *Journal of Finance* 58:2375-2399.

Hendershott, Terrence, and Charles M. Jones. 2005. "Island Goes Dark: Transparency, Fragmentation and Regulation." *Review of Financial Studies* 18: 743-793.

Hillis, Ken, Michael Petit, and Kyle Jarrett. 2013. *Google and the Culture of Search*. New York, Routledge.

Hirschey, Nicholas H. 2011. Do High-Frequency Traders Anticipate Buying and Selling Pressure?, Available at https://www2.bc.edu/~taillard/Seminar_spring_2012_files/Hirschey.pdf, accessed 13 August 2012.

Johnson, Neil, Guannan Zhao, Eric Hunsader, Jing Meng, Amith Ravinder, Spencer Carrran, and Brian Tivnan. 2012. Financial Black Swans Driven by Ultra Fast Machine Ecology, Available at http://arxiv.org/abs/120z.1448, accessed February 21, 2012.

Knight Capital Group, Inc. 2012. 2011 Annual Report, Jersey City, NJ: Knight Capital Group. Available at http://www.knight.com, accessed December 6, 2012.

Knorr Cetina, Karin. 2007. "Global Markets as Global Conversations." *Text and Talk* 27:705-734.

—. 2013. Presentation to Panel, Theorizing Numbers, American Sociological Association Annual Meeting, New York, August 9-13.

Knorr Cetina, Karin, and Urs Bruegger. 2002. "Global Microstructures: The Virtual Societies of Financial Markets." *American Journal of Sociology* 107:905-951.

Knorr Cetina, Karin, and Alex Preda. 2007. "The Temporalization of Financial Markets: From Network to Flow." *Theory, Culture and Society* 24/7-8:116-138.

Krippner, Greta, Mark Granovetter, Fred Block, Nicole Biggart, Tom Beamish, Youtien Hsing, Gillian Hart, Giovanni Arrighi, Margie Mendell, John Hall, Michael Burawoy, Steve Vogel, and Sean O'Riain. 2004. "Polanyi Symposium: A Conversation on Embeddedness." *Socio-Economic Review* 2:109-135.

Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford, Oxford University Press.

Lenglet, Marc. 2011. "Conflicting Codes and Codings: How Algorithmic Trading is Reshaping Financial Regulation." *Theory, Culture & Society* 28:44-66.

Lewis, Michael. 2014. *Flash Boys: Cracking the Money Code*. London: Penguin.

Mackenzie, Adrian. 2006. *Cutting Code: Software and Sociality*. New York: Peter Lang.

Mackenzie, Michael. 2012. "Calls to Extend Circuit Breakers on Superfast Traders." *Financial Times* August 15:26.

Madhavan, Ananth. 2012. Exchange-Traded Funds, Market Structure and the Flash Crash, Available at http://ssrn.com/abstract=1932925, accessed 30 December 2012.

Massoudi, Arash, and Michael Mackenzie. 2013a. "In Search of a Fast Buck." *Financial Times* February 20:11.

—. 2013b. "Investors Turn to the Dark Side for Trading." *Financial Times* April 26: 30.

McInish, Thomas, James Upson, and Robert Wood. 2012. The Flash Crash: Trading Agressiveness, Liquidity Supply, and the Impact of Intermarket Sweep Orders, Available at http://ssrn.com/abstract=1629402, accessed June 3, 2012.

Mehta, Nina. 2011. "Credit Suisse Plans Market for Long-Term Investors." *Bloomberg Businessweek* January 11: Available at http://www.businessweek.com, accessed January 17, 2011.

Muniesa, Fabian. 2003. Des marchés comme algorithmes: Sociologie de la cotation électronique à la Bourse de Paris, PhD thesis: École Nationale Supérieure des Mines.

—. 2005. "Contenir le marché: La transition de la criée à la cotation électronique à la Bourse de Paris." *Sociologie du travail* 47:485-501.

—. 2011. "Is a Stock Exchange a Computer Solution? Explicitness, Algorithms and the Arizona Stock Exchange." *International Journal of Actor-Network Theory and Technological Innovation* 3/1:1-15.

Pardo-Guerra, Juan Pablo 2010a. Computerising Gentlemen: The Automation of the London Stock Exchange, c. 1945-1995, PhD thesis: University of Edinburgh.

—. 2010b. "Creating Flows of Interpersonal Bits: The Automation of the London Stock Exchange, c.1955-90." *Economy and Society* 39:84-109.

Patterson, Scott. 2012. *Dark Pools: High-Speed Traders, A.I. Bandits, and the Threat to the Global Financial System*. New York, Crown.

Patterson, Scott, and Jenny Strasburg. 2012. "Dark Markets: For Superfast Stock Traders, A Way to Jump Ahead in Line." *Wall Street Journal* September 19:A1.

Patterson, Scott, Jenny Strasburg, and Liam Pleven. 2013. "Speedy Traders Exploit Loophole." *Wall Street Journal* May 1: 1-2.

Poon, Martha. 2007. Scorecards as Devices for Consumer Credit: The Case of Fair, Isaac & Company Incorporated. *Market Devices*. M. Callon, Y. Millo and F. Muniesa. Oxford, Blackwell:284-306.

—. 2009. "From New Deal Institutions to Capital Markets: Commercial Consumer Risk Scores and the Making of Subprime Mortgage Finance." *Accounting, Organizations and Society* 34:654-674.

Popper, Nathaniel, and Christopher Leonard. 2012. High-Speed Traders Profit at Expense of Ordinary Investors, a Study Says, *New York Times*, December 3. Available at http://www.nytimes.com, accessed December 5, 2012.

Preda, Alex. 2006. "Socio-Technical Agency in Financial Markets: The Case of the Stock Ticker." *Social Studies of Science* 36:753-782.

—. 2009a. "Brief Encounters: Calculation and the Interaction Order of Anonymous Electronic Markets." *Accounting, Organizations and Society* 34:675-693.

—. 2009b. *Framing Finance: The Boundaries of Markets and Modern Capitalism*. Chicago, University of Chicago Press.

—. 2013. "Tags, Transaction Types and Communication in Online Anonymous Markets." *Socio-Economic Review* 11:31-56.

Saavedra, Serguei, Kathleen Hagerty, and Brian Uzzi. 2011. "Synchronicity, Instant Messaging, and Performance among Financial Traders." *Proceedings of the National Academy of Sciences* 108:5296-5301.

Securities and Exchange Commission. 2005. "17 CFR Parts 200, 201, et al.: Regulation NMS; Final Rule." *Federal Register* 70/124 (June 29):37496-37644.

—. 2013. In the Matter of Knight Capital Americas LLC, Respondent. Available at http://www.sec.gov/litigation/admin/2013/34-70694.pdf, accessed May 7, 2014.

Seligman, Joel. 1982. *The Transformation of Wall Street: A History of the Securities and Exchange Commission and Modern Corporate Finance*. Boston, Houghton Mifflin.

Smith, Robert MacKenzie. 2012. "Has Last Look Had Its Day in FX?" *FX Week* December 10: Available at http://www.fxweek.com/2229300/, accessed October 28, 2013.

Winner, Langdon. 1977. *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought*. Cambridge, MA, MIT Press.

Zaloom, Caitlin. 2006. *Out of the Pits: Trading and Technology from Chicago to London*. Chicago, Chicago University Press.

Zelizer, Viviana A. 2012. "How I became a Relational Economic Sociologist and What Does That Mean?" *Politics & Society* 40(2):145-174.

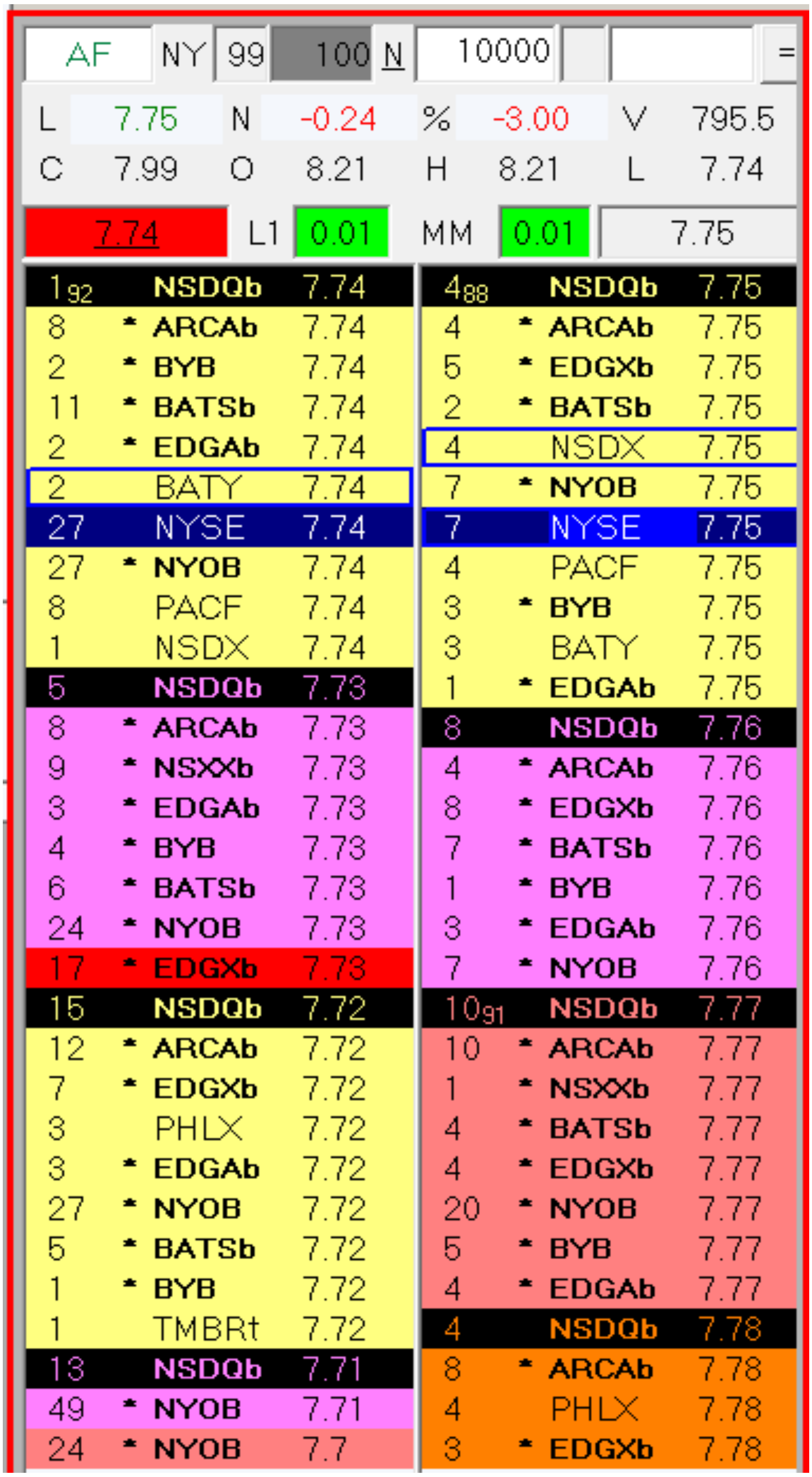
U.S. shares	53%
Global futures	52%
Global foreign exchange	40%
Global fixed income (bonds and bond-like products)	18%

TABLE 1. High-frequency trading as a percentage of all trading in selected markets in 2012. Source: Aite Group estimates; Massoudi and Mackenzie (2013a).

High-frequency traders		43
of which primary area of experience:		
Shares	21	
Futures	11	
Foreign exchange	6	
Other instruments	5	
Exchange/trading venue personnel		44
of which primary area of experience:		
“Lit” share trading	5	
“Dark” share trading	7	
Futures	12	
Foreign exchange	13	
Fixed income	4	
Options	3	
Suppliers/users of execution algorithms		6
Practitioners of other forms of algorithmic trading		8
Manual traders		7
Brokers		3
Suppliers of hardware or services to HFT		12
Market analysts with expertise in HFT		14
Regulator		1
Total		138

TABLE 2. Interviewees.

In total, 125 interviews were conducted, of which 15 were with two people, five with three people, and one with four people. Eight people were interviewed twice, two were interviewed three times, and one interviewed four times. “Lit” and “dark” are explained in footnote 6.



AF | NY | 99 | 100 N | 10000 | =

L 7.75 N -0.24 % -3.00 V 795.5

C 7.99 O 8.21 H 8.21 L 7.74

7.74 L1 0.01 MM 0.01 7.75

1$_{92}$		NSDQb	7.74	4$_{88}$		NSDQb	7.75
8	*	ARCAb	7.74	4	*	ARCAb	7.75
2	*	BYB	7.74	5	*	EDGXb	7.75
11	*	BATSb	7.74	2	*	BATSb	7.75
2	*	EDGAb	7.74	4		NSDX	7.75
2		BATY	7.74	7	*	NYOB	7.75
27		NYSE	7.74	7		NYSE	7.75
27	*	NYOB	7.74	4		PACF	7.75
8		PACF	7.74	3	*	BYB	7.75
1		NSDX	7.74	3		BATY	7.75
5		NSDQb	7.73	1	*	EDGAb	7.75
8	*	ARCAb	7.73	8		NSDQb	7.76
9	*	NSXXb	7.73	4	*	ARCAb	7.76
3	*	EDGAb	7.73	8	*	EDGXb	7.76
4	*	BYB	7.73	7	*	BATSb	7.76
6	*	BATSb	7.73	1	*	BYB	7.76
24	*	NYOB	7.73	3	*	EDGAb	7.76
17	*	EDGXb	7.73	7	*	NYOB	7.76
15		NSDQb	7.72	10$_{91}$		NSDQb	7.77
12	*	ARCAb	7.72	10	*	ARCAb	7.77
7	*	EDGXb	7.72	1	*	NSXXb	7.77
3		PHLX	7.72	4	*	BATSb	7.77
3	*	EDGAb	7.72	4	*	EDGXb	7.77
27	*	NYOB	7.72	20	*	NYOB	7.77
5	*	BATSb	7.72	5	*	BYB	7.77
1	*	BYB	7.72	4	*	EDGAb	7.77
1		TMBRt	7.72	4		NSDQb	7.78
13		NSDQb	7.71	8	*	ARCAb	7.78
49	*	NYOB	7.71	4		PHLX	7.78
24	*	NYOB	7.7	3	*	EDGXb	7.78

FIG 1. Orders for shares of Astoria Financial Corp. on U.S. venues, c. noon, October 21, 2011.
Source: interviewee.

Bids to buy		Offers to sell	
		↑	
		$7.78	400
		$7.77	1091
		$7.76	800
		$7.75	488
192	$7.74		
500	$7.73		
1500	$7.72		
1300	$7.71		
	↓		

FIG 2. Orders for shares of Astoria Financial Corp. on NASDAQ, c. noon, October 21, 2011.

Source: extracted from Figure 1.