

# Systemic failures and organizational risk management in algorithmic trading: Normal accidents and high reliability in financial markets

Social Studies of Science

1–26

© The Author(s) 2021



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/03063127211048515

[journals.sagepub.com/home/sss](https://journals.sagepub.com/home/sss)**Bo Hee Min**  and **Christian Borch**

## Abstract

This article examines algorithmic trading and some key failures and risks associated with it, including so-called algorithmic ‘flash crashes’. Drawing on documentary sources, 189 interviews with market participants, and fieldwork conducted at an algorithmic trading firm, we argue that automated markets are characterized by tight coupling and complex interactions, which render them prone to large-scale technological accidents, according to Perrow’s normal accident theory. We suggest that the implementation of ideas from research into high-reliability organizations offers a way for trading firms to curb some of the technological risk associated with algorithmic trading. Paradoxically, however, certain systemic conditions in markets can allow individual firms’ high-reliability practices to exacerbate market instability, rather than reduce it. We therefore conclude that in order to make automated markets more stable (and curb the impact of failures), it is important to both widely implement reliability-enhancing practices in trading firms and address the systemic risks that follow from the tight coupling and complex interactions of markets.

## Keywords

algorithmic trading, financial regulation, flash crash, high-reliability organizations, normal accidents theory, risk

---

Copenhagen Business School, Frederiksberg, Denmark

### Corresponding author:

Bo Hee Min, Department of Management, Politics and Philosophy, Copenhagen Business School, Porcelaenshaven 18B, Frederiksberg 2000, Denmark.

Email: [bhm.mpp@cbs.dk](mailto:bhm.mpp@cbs.dk)

## Introduction

The days in which the ‘engine rooms’ of financial markets consisted of human traders bustling around so-called trading floors are long gone. During the past few decades, financial markets have undergone an intensive phase of computerization, in which human floor traders have been replaced by automated computer algorithms. While these algorithms are developed by humans, the actual decisions to send orders to buy or sell securities are made by the algorithms themselves and automatically executed through comprehensive technological infrastructure that links individual trading firms to the various exchanges on which they trade. This has dramatically altered the nature of financial markets, as well as the operations of the organizations active on them (MacKenzie, 2018a, 2021; Pardo-Guerra, 2019). In addition to redefining the relationship between humans and technology in the financial industry (Borch and Lange, 2017) and actualizing markets through sociotechnical *agencements* (Callon, 2008), via which human and non-human agents interact (Knorr and Bruegger, 2000, 2002), the rise of algorithmic trading has also introduced new types of market failures and crashes linked to technological risk.

A notable example is the collapse of Knight Capital, a major US trading firm. On August 1, 2012, the firm experienced a shattering algorithmic mishap: A dormant code was unexpectedly triggered, generating millions of erroneous orders, leading to a loss of more than \$460 million in only 45 minutes. Unable to recover from the loss, the firm was soon acquired by a competitor (Kirilenko and Lo, 2013: 65–66). In their subsequent account of the event, the US Securities Exchange Commission (SEC) noted that Knight Capital’s avalanche of orders not only ruined the firm, but also affected other market participants, ‘with some participants receiving less favorable prices than they would have in the absence of these executions and others receiving more favorable prices’ (SEC, 2013: 6). This demonstrates, according to the SEC, that ‘In the absence of appropriate controls [within individual firms], the speed with which automated trading systems enter orders into the marketplace can turn an otherwise manageable error into an extreme event with potentially wide-spread impact’ (p. 2). Indeed, firms must implement a ‘prompt, effective, and risk-mitigating response’ to technological incidents and the ‘failure by, or unwillingness of, a firm to do so can have potentially catastrophic consequences for the firm, its customers, their counterparties, investors and the marketplace’ (p. 3).

The failure of individual trading firms’ algorithmic systems can escalate into wider turmoil because automated markets are deeply interconnected. Since competing algorithmic trading systems digest and respond to the same market data, they make up a highly interactive market ecology, in which one system’s failures and trading decisions can trigger widespread turbulence (Johnson et al., 2013; Sornette and von der Becke, 2011). The most famous example of this is the so-called ‘Flash Crash’ that struck US markets on May 6, 2010. In less than half an hour, market values plunged by \$1 trillion, with the majority of the losses occurring in under five minutes (Menkveld and Yueshen, 2019) and with many market participants quickly withdrawing from the markets (draining liquidity), further exacerbating the crisis. Since then, other flash crash events have occurred, albeit on a smaller scale. However, these events have been frequent enough to

suggest that inter-algorithmic crashes are an ever-imminent problem for present-day automated markets, and represent a new type of systemic risk (Golub et al., 2012; Johnson et al., 2012; World Economic Forum, 2019).

In this article, we analyze the automation of financial markets and their technological risk by revisiting a classical debate about technologically sophisticated organizations and their ability, or lack thereof, to curb their potential contribution, and exposure to failures that might escalate into devastating accidents. Specifically, we argue that Perrow's (1984, 1999) 'normal accident theory' (NAT), with its conception of accident-prone technological systems, and theorization around 'high-reliability organizations' (HROs), with its emphasis on the capacity of certain organizations to curtail technological risk (e.g. Roberts, 1990; Weick and Sutcliffe, 2015), are analytically relevant when it comes to understanding automated markets and the individual organizations that are active on them. The NAT and HRO perspectives were developed prior to the rise of new challenges to organizational operations, including those that have emerged from the ubiquity of digital technologies. In addition, these streams of research antedate the growing calls for studies on the interactions between an adverse environment and organizations and on how organizations build resilience to effectively prepare for, respond to, and mitigate adversity (e.g. Gephart et al., 2009; Ramanujam, 2003; van der Vegt et al., 2015; Williams and Shepherd, 2016; Williams et al., 2017). This includes work on financial markets inspired by science and technology studies (STS). For example, Beunza (2019) examines risk management in financial markets with a particular emphasis on the ways various economic models are deployed to deal with uncertainty (similarly Hansen and Borch, 2021). Despite this recent literature, we nonetheless focus on the classical NAT and HRO perspectives because they highlight important concepts that encompass the later scholarship on organizational risk and resilience. For example, reliability represents 'an intersection of effectiveness, safety, and resilience' (Carroll, 2018: 37), just as NAT prefigures central elements from later debates around organizational risk (Gephart et al., 2009). Indeed, the NAT and HRO perspectives capture the tensions and connections between overall systemic risk and failure in technology-rich settings and firm-level organizational behavior, including risk-management practices. We believe these tensions and connections are particularly salient when it comes to understanding automated trading and its risks.

We argue that present-day automated markets exhibit characteristics that are associated with normal accidents. In line with the HRO scholarship, we further argue that, in spite (or because) of these NAT characteristics, there are good reasons to endorse HRO principles in specific trading firms, in part to avoid collapses like that of Knight Capital, and in part to prevent failures that trigger wider market avalanches. However, we also argue that automated markets are characterized by features that are underappreciated in both NAT and HRO scholarship. Each of these traditions tend to treat their units of analysis (technological systems and organizations, respectively) as insulated entities that have no close interaction with their environments. This analytical problem is particularly relevant and consequential for HRO research, as it means that even a widespread implementation of HRO principles might be deficient when it comes to preventing large-scale accidents in automated markets, where the level of interconnection between algorithms and environments is towering. Indeed, we argue that, paradoxically, particular conditions

and ways of organizing markets might mean that the implementation of high-reliability practices in individual organizations contributes to systemic accidents. Consequently, we propose that, while there are good reasons to promote HRO principles in this industry, these should be combined with measures, including regulatory ones, that tie organization-internal dimensions to the systemic, extra-organizational aspects of present-day financial markets. In other words, it is important to consider organizational and market-wide dimensions simultaneously.

Our analysis contributes to the discussions in STS regarding algorithmic trading. Extending the broader STS literature on financial markets (e.g. Beunza, 2019; Callon, 1998; MacKenzie, 2006; MacKenzie and Spears, 2014a, 2014b; Pinch and Swedberg, 2008), scholars have studied the ways in which market automation is embedded in complex technological systems. For example, MacKenzie has demonstrated that algorithmic trading in the form of so-called high-frequency trading (consisting of fully automated, high-speed algorithms) relies on particular material infrastructures, such as microwave and fiber-optic transmission systems (MacKenzie, 2015b, 2017a, 2017b, 2018b, 2021). Similarly, Pardo-Guerra (2019) has detailed the workings of today's automated financial exchanges as well as their historical backdrop. Other researchers have addressed various organizational dimensions, including trading firms' concerns with algorithmic debugging (Seyfert, 2016), the types of organizational ignorance that pertain to algorithmic trading (Beverungen and Lange, 2018; Lange, 2016; Souleles, 2019), regulatory efforts to change the behavior of trading firms (Coombs, 2016), and conflicting intra-organizational conceptions of algorithms (Lenglet, 2011). We add to this literature by analyzing how the technological risk of market automation is linked to particular intra- and extra-organizational practices.

By focusing on algorithmic trading, this article also advances NAT and HRO debates, as neither tradition has achieved a strong resonance in discussions about financial markets. The first edition of Perrow's *Normal Accidents*, published in 1984, does not address financial markets. In the afterword to the 1999 edition, Perrow (1999: 385–387) briefly discusses 'financial systems' – mentioning 'high-speed trading' in passing – without offering a systematic examination. NAT has been applied in a few studies regarding the 2008 financial crisis (Curran, 2020; Guillén and Suárez, 2010; Palmer and Maher, 2010; Schneiberg and Bartley, 2010), although Perrow (2010) contests that this crisis was a true normal accident, arguing instead that it was caused by human actions, such as financiers' wrongdoing. NAT has seen only few applications in examinations of algorithmic trading (for a conceptual discussion of normal accidents in AI-based automated systems, see Maas, 2018), though Lange et al. (2019: 609) briefly note that automated markets exhibit NAT characteristics. We use the Flash Crash as a way of fleshing out the NAT features of these markets.

While there have been some applications of NAT vocabulary in debates about financial markets, HRO notions are virtually nonexistent in the analyses of this field. The few HRO studies that do attend to the financial industry primarily focus on the absence of high-reliability principles (Bush et al., 2012; Weick and Sutcliffe, 2015; Young, 2012), with only one brief, early study identifying a commitment to high-reliability procedures (Roberts and Libuser, 1993). The relative absence of any discussion of the financial industry in the HRO literature is surprising. Financial markets are extremely risky

environments, prone to recurrent crises and crashes, but some organizations manage to operate in a reliable manner in situations where reliable operations are obviously demanded by regulators.

## **Organizational approaches to failures in technological systems: Revisiting the NAT-HRO debate**

NAT and HRO each constitute a dominant approach to understanding critical failure in organizational theory (Leveson et al., 2009; Rijpma, 1997, 2003; Sagan, 1993; Shrivastava et al., 2009; Vaughan, 2005). The two schools of thought have made significant contributions to our understanding of the performance of organizations in highly technical environments that involve both humans and technology. Although leading representatives of the two approaches point out that the two theories are not necessarily contradictory (LaPorte, 1994; Rijpma, 1997, 2003; Vaughan, 2005), there has been a longstanding debate about whether the perspectives they provide complement or contradict each other (Shrivastava et al., 2009: 1357).

NAT's emphasis is on a negative extreme of organizational performance – a catastrophic accident that affects a system in its entirety, which Perrow refers to as a normal or system accident (Perrow, 1981, 1984, 1999). Working from the case study of the near-meltdown accident at the Three Mile Island nuclear power plant, Perrow argues that risk is inherent in technology and accidents *will* happen no matter what organizations do – often in the absence of any human error. According to Perrow, two factors render technological systems susceptible to normal accidents. The first is complex interactions, which refer to 'unfamiliar sequences, or unplanned and unexpected sequences' between systems components that are 'either not visible or not immediately comprehensible' (Perrow, 1999: 78). Complex interactions may involve feedback loops and are contrasted with linear interactions, whose sequences are expected and comprehensible. The second factor is tight coupling between components in the system, via which the output of one is the input of the other, with little flexibility. Tight coupling leaves humans insufficient time or little possibility to intervene. Perrow (1999) also argues that safety measures implemented in the design or construction of the system, including engineered safety features or redundant parts – such as the secondary O-ring in the Challenger launch accident (Vaughan, 1996) – have a limited role in preventing accidents. These safety measures, on the contrary, are said to often increase the risk of accidents by increasing complexity, creating unanticipated interactions, or inducing overconfidence. As Vaughan concludes in her analysis of the Challenger accident, 'While good management and organizational design may reduce accidents in certain systems, they can *never* prevent them' (Vaughan, 1996: 416, emphasis added).

HRO scholars, on the other hand, examine the organization as the unit of analysis and, in contrast to the pessimism and technological determinism of NAT, theorize the alternative extreme of organizational performance – organizations can attain high-level reliability through the successful prevention of failures over an extended period. HRO theorization suggests that technologically rooted risk can be effectively managed by good organizational design and practices. This constructs accidents as the result of suboptimal or poor management. For empirical cases, HRO scholars have studied organizations that have a

long history of reliability while engaging in high-risk activities, such as aircraft carriers, air traffic controls, and firefighters. Building on such studies, Weick and Sutcliffe (2001) argue that HROs are characterized by five dimensions: (1) a preoccupation with failure identification and reporting, (2) a reluctance to simplify interpretations of incidents, (3) a sensitivity to the organization's operational aspects, (4) a commitment to resilience, and (5) a deference to expertise (rather than hierarchy) when dealing with incidents. These features help to ensure that HROs develop a culture (with corresponding incentives) that is committed to detecting, reporting, acting upon, and learning from small incidents before they escalate into large-scale accidents (see LaPorte and Consolini, 1991; Roberts, 1989, 1990; Sagan, 1993; Weick, 1987; Weick and Sutcliffe, 2015).

Although both perspectives on organizational failure provide useful concepts and a better understanding of technological risk, they have limited application when it comes to many contemporary organizations. NAT suggests that, where the risks outweigh the benefits for a particular technology, we should replace it with a safer technology to prevent accidents. For example, regarding nuclear power plants, Perrow (1999: 348) argues, 'the case for shutting down all nuclear plants in the United States seems to be clear'. Retiring all existing systems built on risky technology, however, may not always be possible or necessary if we can effectively manage them (Roberts, 1993: 166). Similarly, HRO theorization faces some limitations. Because the theory is based on institutions that are known for their low accident rates, HROs are unlike the majority of organizations. Most of the organizations praised by HRO scholars are public institutions, which are often monopolistic providers of given functions and rarely face competitive pressure, and an emphasis on reliability is therefore not challenged by imperatives of competition and shareholders' desire for higher profit.

Moreover, both NAT and HRO analyze organizations as insulated from external environments, including other organizations. For example, Perrow (1999) focuses on 'properties of systems' (p. 63), and although such systems can encompass many scales, he does not systematically investigate how interactions and couplings might exist not only inside high-risk systems, but also between such systems and their environments. Perrow (1999: 75) does observe that, regardless of their internal properties, practically all types of technological systems 'will have at least one source of complex interactions, the environment, since it impinges upon many parts or units in the system'. However, despite such gestures, he does not analyze the role of the environment in a systematic manner. Perrow's (1999) analyses primarily treat the environment in terms of natural phenomena such as bad weather or falling rocks, which constitute challenging conditions for marine officers and miners, respectively (pp. 176, 251). The main exception is his discussion of military systems that seek to detect and respond to missile attacks, in which the environment is intentional and self-activating (p. 292). Early-warning and response systems are intimately linked to their environments and enemy systems in that they are operating under the same conditions – they, too, are constantly monitoring their environments and responding to any warnings they detect. That said, Perrow's discussion of such highly interactive system–environment configurations is brief and not elaborated upon theoretically.

Likewise, the richness with which HRO scholars portray the inner workings of HROs is not paralleled by a similar interest in the environmental contexts in which these organizations are situated (Roberts, 2018: 6). Weick and Sutcliffe (2007: 85) merely note that a

mindful HRO approach is particularly salient ‘in contexts that are dynamic, ill structured, ambiguous, or unpredictable’. Roberts similarly observes that HROs ‘face very uncertain environments’ and mentions as an example that aircraft carriers may be exposed to unexpected weather conditions (Roberts, 1990: 161, 71; see also Bigley and Roberts, 2001; Sutcliffe, 2011). The most elaborate HRO discussion concerning environmental factors stems from Roe and Schulman (2016), whose central concern is how infrastructures such as telecommunication networks or power grids are interconnected such that the failure of one infrastructure might adversely affect another. While they mention important organizational aspects related to ensuring the high reliability of interconnected infrastructures, their analysis does not directly address the more general question of how organizations are linked to their environments and the HRO implications of such links.

Furthermore, the ubiquity of information and communication technologies has dramatically increased the connectivity and interdependence of contemporary organizations to environments and other organizations, making it imperative to manage exogenous risk while ensuring endogenous reliability. The technological systems within organizations are complex and interconnected, and these organizational systems connect to one another to create a large-scale system or system of systems, increasing the opportunity for a failure to spread its disruptive influence rapidly system-wide (Cliff and Northrop, 2012; Helbing, 2013).

## Data and methods

Our analysis is based on a broader examination of the algorithmic trading industry. Working in collaboration with colleagues (Kristian Bondo Hansen, Nicholas Skar-Gislinge, Pankaj Kumar, and Daniel Souleles), between September 2017 and December 2020, we conducted 189 semi-structured interviews with financial market participants at 141 institutions (77 of these interviews were conducted collaboratively or independently by the authors). The majority of these interviews were in firms specializing in algorithmic trading, such as proprietary trading firms, banks, and hedge funds. Most of these firms were located in Chicago, New York, London, or Amsterdam. We also interviewed brokers, exchange officials, regulators, technology providers, and institutional investors. In the interviews, which typically lasted approximately one hour, though some were significantly longer, we asked questions that helped us shed light on how algorithmic trading is reshaping financial markets and the firms active on them.

Our analysis is informed by this broader investigation, but it particularly draws on data collected at Tyler Capital Limited, a London-based firm specializing in algorithmic trading centered on machine-learning (ML) models. Established in 2003, Tyler Capital is a proprietary trading firm (trading on its own account rather than on behalf of clients). With a staff of approximately 50, Tyler Capital is a medium-sized firm. Its focus is on high-frequency trading in futures contracts on the Chicago Mercantile Exchange (CME), although it is also active in other markets around the globe. Since 2014, the firm has focused exclusively on ML-based trading. Instead of human traders developing trading strategies that are then coded into and implemented by the algorithmic system, it is the firm’s ML system that comes up with trading strategies based on the data it is fed.

Despite operating in an industry that is notoriously secretive, Tyler Capital granted us rare access to its work. Our case data consist of interviews, internal documents, and ethnographic observations. Interviews are an efficient way to collect rich empirical data for episodic and unprecedented phenomena such as unexpected disruptions and incidents in which high-reliability practices are most visible (Eisenhardt and Graebner, 2007: 28). Since November 2017, we have conducted 23 formal, semi-structured interviews with 18 people from Tyler Capital, resulting in 429 pages of transcripts. We interviewed the founder and the management team, and conducted several interviews with the Chief Executive Officer (CEO) and Chief Technology Officer (CTO), in particular, since they had been instrumental in transforming the firm into one that emphasizes ML and high reliability. We interviewed several other members of the firm more than once, as they were highly knowledgeable informants about technological risk. In order to gather data on practices throughout the entire organization, we interviewed people from all functional and cross-functional teams and at all hierarchical levels. Also, we spent most of a day off-site with the CTO discussing related topics.

In addition, we made three ethnographic visits to the firm over the course of one-and-a-half years, each lasting a couple of days. On our visits, the firm provided us with access cards and the opportunity to observe all parts of the organization. We were able to have discussions in which employees explained their work and tools, and to engage in participant and non-participant observation in management meetings and information-sharing meetings. Finally, the management granted us access to confidential internal documents, totaling 324 pages, which outlined numerous financial, strategic, organizational, and operational aspects of the firm, including incident reports and protocols for dealing with potential incidents.

The different data sources allowed us to triangulate our findings, thereby ensuring validity. Our repeated interactions with individuals from all parts of the organization permitted us to pursue specific themes of relevance to technological risk over longer periods of time, and to request further data (e.g. internal documents) where necessary. All such requests were met by management. Among other things, this allowed us to explore what Tyler Capital's management and staff would regard as 'incidents', focusing on individual members' responses and the coordination thereof. As an interview with any single informant would hardly provide a complete picture of the organization-level response to an incident, we tracked relevant members' activities during an incident. We also conducted a within-case analysis of the responses in relation to the envisaged repercussions of any unaddressed incidents. Further validity was ensured by comparing the case data to our larger examination of the algorithmic trading industry (the pool of 189 interviews).

Although we only discuss a few concrete examples of Tyler Capital's risk-management emphasis in this article, it is worth stressing that we identified several additional examples during our fieldwork. These include procedures to avoid data issues that might negatively affect the ML system's learning and subsequent trading decisions, as well as organizational measures to ensure that the system does not engage in manipulative behaviors.

The management's decision to waive the firm's anonymity for this study is rare for this type of work but not without precedents, as other studies of algorithmic trading exist

in which some informants have similarly chosen to be identified (e.g. MacKenzie, 2017b, 2021). Tyler Capital's management may have had several reasons for waiving anonymity. Through our conversations with the management and staff, it became clear to us that they consider their firm as occupying an exceptional position in the markets, including being at the forefront of ML-based trading, and strongly emphasizing market integrity, as well as seeing themselves as leaders in the domain of risk management and reliable operations and hoping to set the standard for the industry on these matters. Waiving anonymity might be one way to achieve this. Although waiving anonymity could be said to potentially constrain our analysis (making a critical discussion of the firm's work less comfortable for us), we did not experience any actual constraints on our access to data and analytical work, nor did we meet any manifest or tacit pressure to present a rose-tinted picture of the firm. For example, as our discussion demonstrates, even though Tyler Capital's management and staff are explicitly committed to HRO principles and have tried to implement these widely in the organization, the firm has encountered incidents – and the management and staff shared detailed information about these with us – which suggests that, despite its efforts, the firm's internal risk-management procedures may not fully anticipate and prevent disasters from ensuing.

Are the findings from Tyler Capital generalizable? On the one hand, comparing the data from Tyler Capital with our broader qualitative examination of the automated trading industry suggests that the firm is a special case on certain dimensions and that it has likely gone further than most other algorithmic trading firms when it comes to implementing HRO principles throughout the organization. On the other hand, it is not uncommon that algorithmic trading is conducted by smaller proprietary trading firms where the owners' own personal capital is at stake, meaning that such firms are subject to strict oversight of the owners and naturally incentivized to implement high-reliability practices of some shape or form (including developing dependable software). Further, regardless of their specific strategies, all algorithmic trading firms operate in the same overall technology-infused market context. Given this, we believe that our discussion of Tyler Capital is helpful in shining a light on some of the types of risk that different algorithmic trading firms are similarly exposed to.

In addition to the data already mentioned, our discussion of the 2010 Flash Crash also draws upon documents from US government agencies, including congressional hearing transcripts, reports from regulators, and minutes and transcripts of meetings and hearings from regulators. These reports largely addressed aggregate- and market-level activities. The official documents were collected primarily from three agencies: the Government Publishing Office, the SEC, and the Commodity Futures Trading Commission (CFTC).

## **Failures and reliability in algorithmic trading**

To analyze technological failures in algorithmic trading, we demonstrate the salience of NAT and HRO by examining market-wide failures and firm-level failures. In doing so, we analyze what we perceive to be the properties of normal accident-prone systems and the characteristics of HROs in financial markets as a large-scale complex system.

### *Normal accidents in algorithmic trading*

The two central features that Perrow associates with normal accidents – complex interactions and tight coupling – are fundamental to automated financial markets. The overall organization of contemporary financial markets makes them rife with complex interactions: Financial markets are largely fragmented by numerous trading venues serving as marketplaces, such as the thirteen national stock exchanges and more than 40 alternative trading venues in the US. A vast number of market participants connect to these to retrieve market information and interact with other participants, each of them pursuing individual strategies and each endowed with particular time horizons, capital, and so on. The complex interactions are also an effect of the fact that trading venues in the US interact with one another to aggregate market information into a single national market, making the automated financial markets a large-scale system. Tight coupling arises because trading firms and brokers typically design their algorithmic trading and trade-execution systems such that they consider the action of other market participants, including other algorithms. Algorithms do so by monitoring the so-called electronic order book, which records all orders sent by market participants to buy or sell securities at a particular exchange. Trading firms may, for example, deploy algorithms that seek to make a profit by detecting a larger market move: If a large number of orders to buy a particular security are quickly piling up in the order book, this suggests that a price increase is imminent; the algorithm may then take advantage of the information by rapidly buying the security and selling it when the price has increased. This simple algorithmic strategy may, due to its simplicity, no longer be widely profitable (MacKenzie, 2019). Nonetheless, it is illustrative of the basic approach of trading algorithms (MacKenzie, 2018a). Given that the fastest algorithmic strategies can respond to order-book changes in a matter of micro- or even nanoseconds (millionths and billionths of a second, respectively), this tight coupling through automated responsiveness in effect leaves little room for human intervention.

Tight coupling manifests itself in the ways in which markets and exchanges are connected, in the sense that failures in one part can pervade the entire system. For example, the same securities may be traded on multiple exchanges simultaneously, and through high-speed algorithmic arbitrage strategies, price changes for a security traded on one exchange will quickly lead to parallel price changes for the same security traded on other exchanges. Similarly, different types of securities are also connected in ways that lead to tight coupling. For example, as detailed by MacKenzie (2017a, 2017b, 2018b), for futures contracts on underlying stock indices, price changes in the former tend to trigger price changes in the latter. This happens in a timeframe of milliseconds, namely the time it takes to transmit information (via microwave or fiber-optics) from trading firms' computer servers, placed in physical proximity to the data center of the CME in Aurora, Illinois, to the New Jersey-based data centers of the New York Stock Exchange (NYSE) and Nasdaq. Accordingly, automated financial markets constitute a highly intensified version of Perrow's early-warning and response systems. While markets are not military systems, market participants see their competitors as 'intentional and self-activating', and know that they have a similar view of trading organizations in *their* environments.

While machine learning-based algorithms are gaining traction within financial markets (Hansen, 2020, 2021; Hansen and Borch, 2021), most algorithms used by brokers and trading firms continue to carry out human-defined instructions in response to particular market movements. This might suggest that algorithmic interactions in markets take place in a predictable, linear fashion, with each algorithm responding in a predictable way to others' actions in markets. However, algorithmic interaction patterns are often nonlinear and unpredictable (Borch, 2020; MacKenzie, 2019), showing characteristics of the complex interactions that are associated with normal accident-prone systems. Indeed, financial markets can be seen as a large-scale complex system composed of individual trading firms' systems. Since the latter are themselves complex systems, comprised of subsystems and units, markets exhibit features of 'hyper-connectivity' and 'hyper-risks' (Helbing, 2013). This is why failures of individual trading firms (such as Knight Capital) can have adverse effects on markets more broadly.

The 2010 Flash Crash remains the most notable example of a market-wide failure in algorithmic trading. The event has drawn much attention, and several explanations have been proposed. Initially, market participants pointed to possible trading errors, so-called fat finger trades, but this type of explanation was soon replaced by accounts that favored a more systemic perspective that aligns the event with NAT features. The CFTC and the SEC, which preside the US futures and stock markets, respectively, offered extensive accounts of the event. According to their official report published five months after the event, the Flash Crash was a result of a confluence of events that *accidentally* developed into a plunge in the futures market, which subsequently spread to the stock market. A large sell order of E-Mini S&P 500 futures contracts by a Kansas City-based mutual fund was executed by an automated algorithm until the fulfillment of the sale of all 75,000 futures contracts (worth \$4.1 billion). This triggered a feedback loop of selloffs among high-speed algorithms that predicted that the sell pressure in the order book was indicative of further price decline. As more algorithms joined the selloff in response to these order book changes – more than 15,000 traders, pursuing different kinds of strategies, were trading the E-Mini during the event (Kirilenko et al., 2017) – the prices spiraled down to unprecedented levels in a matter of minutes and even reached the lowest possible price (a penny) for some stocks (CFTC–SEC, 2010; US House, 2010; US Senate, 2010).

Simultaneously, the computer systems on the NYSE experienced a failure that created delays of up to multiple seconds. This caused concerns among market participants, who questioned the accuracy of the received information (we discuss this in more detail later). Faced with potentially erroneous data, many trading algorithms halted trading as a pre-programmed fail-safe (CFTC–SEC, 2010: 35). Meanwhile, due to the NYSE's technological failure, many orders were re-routed to Nasdaq, whose systems were overwhelmed by the surge in the order volume (CME Group, 2010). This added confusion to both algorithms and human traders, who changed their choice of trading venue or paused trading altogether, further exacerbating the price decline. Both the futures and stock markets rebounded after the CME issued a five-second halt in all trading, whereas the trading halts activated on the NYSE prior to it had no effect on mitigating market distress.

As this demonstrates, the progression of the Flash Crash exhibits the properties of a normal accident – complex interactions among a high number of market participants and

tight coupling among and across these trading venues. In addition, as Perrow argues, fail-safes such as pre-programmed trading pauses in algorithms and trading halts ('circuit breakers') on trading venues did not prevent local failures from bringing down the entire financial system. Rather, they created new interactions that exacerbated market disruptions.

The 2010 Flash Crash was followed by similar subsequent events. For example, in October 2014, the US Treasury securities market suffered a flash crash, wherein the prices of the Treasury notes soared for six minutes and then plunged right back down in the following six minutes with no apparent reason (Department of Treasury et al., 2015). These types of events have inspired a host of research into automated markets and the types of technological risks associated with them. A central point in much of this research is that flash crash events, triggered by complex interactions and tight coupling, are a normal occurrence in present-day markets. For example, based on US market data, Johnson et al. (2012, 2013) and Golub et al. (2012) suggest that, on average, approximately fourteen smaller flash crashes occur every trading day (Borch, 2016). Sornette and von der Becke (2011: 3) further argue that, 'As a consequence of the increasing inter-dependences between various financial instruments and asset classes, one can expect in the future more flash crashes involving additional markets and instruments.'

We have argued that algorithmic trading systems are designed to monitor and potentially respond to any changes in the electronic order book such that the trading of one type of security at one exchange can lead to changes at other exchanges and/or the trading of other securities. In addition to such tight coupling, automated markets consist of numerous market participants and trading venues with a host of complex interactions (with unanticipated effects) playing out among them. From a NAT perspective, this endows automated markets with technological risk, rendering them prone to accidents, be they of a firm-level character (as with Knight Capital-esque collapses that have potentially larger ripple effects on markets) or flash crash types of accidents, stretching from smaller events that go largely unnoticed to events that exercise a huge impact across markets.

### *HRO and algorithmic trading*

In response to market crashes, regulators have taken steps to enhance the stability of markets by promoting a series of measures that curtail some of their complex interactions and tight coupling. For example, the 2010 Flash Crash prompted US regulators to strengthen circuit breakers, which automatically pause trading if prices are moving too much too quickly. In addition, the SEC enacted a new regulation called Regulation Systems Compliance and Integrity in 2014 'to strengthen the technology infrastructure of the US securities markets' as a post-hoc measure of such technology-related market disruptions (SEC, 2014). Similarly, in 2018, a new legislative framework, the 'Markets in Financial Instruments Directive (recast)' (MiFID II), took effect in Europe. MiFID II explicitly addresses algorithmic trading:

An investment firm that engages in algorithmic trading shall have in place effective systems and risk controls suitable to the business it operates to ensure that its trading systems are resilient and have sufficient capacity, are subject to appropriate trading thresholds and limits

and prevent the sending of erroneous orders or the systems otherwise functioning in a way that may create or contribute to a disorderly market. Such a firm shall also have in place effective systems and risk controls to ensure the trading systems cannot be used for any purpose that is contrary to Regulation (EU) No 596/2014 or to the rules of a trading venue to which it is connected. The investment firm shall have in place effective business continuity arrangements to deal with any failure of its trading systems and shall ensure its systems are fully tested and properly monitored to ensure that they meet the requirements laid down in this paragraph. (European Parliament and Council of the European Union, 2014: Article 17, 1)

MiFID II does not specify what constitutes ‘suitable’ risk controls, ‘effective business continuity arrangements’, or proper monitoring. Similarly, MiFID II requires an ‘appropriate testing of algorithms’ (Article 48, 6), including ‘appropriate stress testing’ (Article 9, 3b), without this being further delineated. Our fieldwork suggests that, on the one hand, trading firms would generally be committed to several of these requirements. For example, it would be a standard procedure for the firms we interviewed to perform comprehensive backtesting of their strategies, just as firms would tend to scale their strategies only gradually. Both measures serve the purpose of minimizing a firm’s risk exposure when launching algorithmic strategies that are part of markets characterized by complex interactions and tight coupling. Similarly, firms would have ‘kill switches’ implemented, which can be activated if markets come under significant stress (as during the Flash Crash), and which would immediately cancel all of the firm’s outstanding orders.

On the other hand, although it seems reasonable to say that MiFID II’s underlying objective is to improve the stability of automated markets through reliability-enhancing measures, the regulation leaves much leeway for trading firms to define their own approach. Accordingly, it is conceivable that a trading firm might formally comply with MiFID II without it actually considerably enhancing market stability. We suggest that a proper commitment to HRO principles in trading firms is a way of achieving, formally and substantially, the regulatory objectives set out in MiFID II and similar pieces of regulation.

To illustrate how this might materialize, we now turn to our fieldwork at Tyler Capital. A London-based firm, Tyler Capital must comply with MiFID II (this remains the case post-Brexit). Similar to other algorithmic trading firms, a critical element in Tyler Capital’s technological infrastructure consists of its connectivity to markets. Connectivity can have more dimensions, but in particular, it concerns how algorithmic trading firms are hooked up with exchanges. Through third-party network connections to exchanges, Tyler Capital can observe the electronic order book via real-time market data, and closely monitor and manage its ML trading system and the firm’s risk profile, including outstanding orders to buy or sell securities. Most algorithmic trading firms rely on third-party network providers for such connectivity. The network provider used by Tyler Capital operates transatlantic submarine cables that connect London to New York, and then to Chicago. Specifically, to ensure redundancy, the network provider operates two transatlantic high-speed cables, which are deliberately placed at a considerable distance from each other.

However, one day in February 2017, an improbable event caused an outage of the high-speed network: A ship dragging its anchor severed both of the cables. Tyler Capital’s

*trading* system was unaffected by this, as it is ‘co-located’ in the CME’s data center in Aurora, Illinois. In other words, the firm has a server placed in physical proximity to the exchange’s computer system. However, although actual trading could continue despite the outage, Tyler Capital lost its high-speed *system management and monitoring capacity*. As a result, the firm was unable to closely monitor the orders to buy and sell securities that the ML system had placed at the exchange.

Faced with this externally generated disruption, Tyler Capital activated a series of measures to ensure reliability. Upon receiving an automated alert from its technology infrastructure system that monitoring capacity had been lost, the firm initiated a pre-defined incident-management procedure, led by the head of trading. Collaborating with the infrastructure engineers at the firm and external partners such as the network vendor and the CME, the head of trading activated a manual market-monitoring tool provided by the CME, and assessed the root cause and business impact of the incident. Concluding that manual monitoring of the ML trading system was insufficient to ensure the highest level of reliability, he activated the kill switch. An organization-wide process to address this unanticipated event – which involved the reporting of the incident, the convention of key members of the firm, an assessment of the associated risk, the evaluation and selection of post-hoc measures, and the implementation of the selected measure – was carried out within 30 minutes of the firm becoming aware of the incident. While the firm had an alternative means to monitor its own trading activities, it completely withdrew from the affected market, because that alternative was slightly slower than the lost connection and thereby would not warrant sufficient reliability.

The decisions and actions taken were recorded in the firm’s incident report. When the network service was restored approximately twelve hours later, all relevant members of the firm carried out an extensive system check, along with an ‘extra-vigilant reconciliation’ of all orders, and a backup network ran for another twelve hours from the point of network recovery (cited from internal documents). The members of the firm also collectively conducted a post-mortem analysis to locate the root cause of the incident and assess its business impact, as well as to evaluate the adequacy and responsiveness of their incident management.

This example encapsulates the defining characteristics of ‘classical’ HROs: a preoccupation with failure (vigilant monitoring and response for even non-critical incidents), reluctance to simplify interpretations (a network outage can be more than just a loss of connectivity), sensitivity to operations (encompassing all operations, including network connectivity and the trading system), commitment to resilience (post-mortem analysis and checkup), and deference to expertise (the head of trading rather than the firm’s ML experts led the incident-management process). The example also illustrates the risk exposure of organizations that are tightly coupled to their environment. While the incident was in one sense non-critical, because the firm’s actual trading activities could continue unaffected, it was deemed critical because it originated in the environment, and hence beyond the organization’s sphere of control, which compromised the reliability of the ML trading system. Because of its reliance on a third-party service over which it definitionally had no operational purview, Tyler Capital felt particularly vulnerable. In fact, they reasoned that, in this particular situation, the best way to ensure high reliability was to activate the kill switch and thereby sever the firm’s link to the market. Crucially, while

it is routine for algorithmic trading firms to press the kill switch if they face excessive losses, Tyler Capital exercised this option when faced only with a loss of monitoring capability.

The connectivity example demonstrates how a commitment to suitable risk controls, effective business continuity arrangements, and proper monitoring – all demanded by MiFID II – may be implemented in trading firms. As mentioned earlier, trading firms are also required to stress test their algorithms so as to reduce both their exposure and their potential contribution to algorithmically generated market failures and to ensure that the algorithms continue to work effectively under adverse market conditions. Instead of pursuing a minimalistic approach that might ensure *pro forma* compliance, Tyler Capital’s commitment to reliability and resilience has led it to expose its ML trading system to extreme stress testing. In the words of the CTO, ‘We’re not saying, “This is our minimum requirement”. We’re saying, “What’s the right thing to do? Let’s codify it”. To that end, in the fall of 2018, the company hired a leading US law firm specializing in financial regulation and litigation to run a simulated enforcement exercise. The aim was to examine the robustness firstly of the ML trading system when faced with an extreme market event, and secondly of the firm’s ability to account for its operations and procedures vis-à-vis the types of penetrating questions that regulators would typically ask in such situations. In other words, the exercise included both a simulated market-disruption event and a subsequent mock litigation.

In practice, Tyler Capital ‘created a scenario to say, “Let’s try and find an example where something could go catastrophically wrong and let’s then examine all of our processes and our approach and our methodologies in light of that” (senior software engineer). Aside from its catastrophic nature, it was deemed crucial that the simulated event had some degree of realism from a market point of view. Specifically, the scenario included a combination of failures that made the different strategies of Tyler Capital’s trading system interact in a devastating manner, creating large and aggressive positions that eventually unleashed a market panic and exposed the firm to massive risk, all of which occurred while its usual monitoring of the ML system was defective. Importantly, the scenario demanded that Tyler Capital lessens some of its controls: ‘we needed to make the [trading] system do some things that it’s not actually allowed to do, in order to create that particular scenario’ (senior software engineer).

From an HRO point of view, what is important about this exercise is that, while Tyler Capital was reassured that its system and procedures were highly robust – the law firm concluded that ‘the protections that currently surround the operation of the [ML trading system] are extremely strong’ (cited from internal documents) – it also used the exercise as a learning experience to reflect on whether particular aspects of its work and ways of organizing could be improved. For example, some monitoring procedures were formalized and strengthened.

We emphasize these examples from Tyler Capital as an illustration of how algorithmic trading firms may seek to implement HRO principles in order to minimize the risk of technological failures and incidents escalating into large-scale accidents. In addition, the examples make it clear that, in the context of automated markets, HRO implementation needs to address the fact that algorithmic trading firms are deeply interconnected with their environment. This speaks to our earlier remarks about the central analytical

drawback of existing HRO scholarship: It has a proclivity to treat organizations as insulated entities that may achieve high levels of reliability by vigorously detecting and responding to any *internal* incidents – rather than analyzing organizations in the context of the environments in which they are embedded and to which they may have important connections. In automated markets – and more generally, in many data-dependent contexts – firms are also exposed to incidents in their environment, just as any internal incidents may quickly escalate outward and adversely affect others. What emerges from this is the need for a new and encompassing risk appreciation, according to which an HRO should consider itself accountable for any externally induced adverse effects, as they indicate a lack of internal due diligence and reliability. As stated by van der Veegt et al. (2015: 972), an HRO conceived as such entails a reorientation from traditional risk management to a resilience approach, in which a ‘preparedness for dealing with unforeseen disruptive events’, be they internally or externally generated, is key.

## Discussion

Compared with the pessimism guiding NAT, our HRO analysis is a little more optimistic, in that we would attribute HRO practices an important role in reducing the negative effects of complex interactions and tight coupling. In that sense, our analysis is akin to Coombs’s (2016) discussion of regulatory attempts to reduce the systemic risks of algorithmic trading. Coombs focuses on the 2013 German High-Frequency Trading Act, which introduced a so-called algorithm-tagging rule that enabled regulators to return to trading firms and identify which algorithms generated particular trading decisions. He finds that, although regulators were well aware that the Act was imperfect, it had a series of positive effects, including empowering compliance officers in trading firms, in effect endowing them with greater powers and integrating them better into the operational layers of their respective firms. Coombs appreciates that his positive analysis might come across as ‘rose-tinted’ and that the regulators he interviewed might have had an overly optimistic view of their own work (p. 294). That said, he cautions against what he sees as an automatic reaction in the sociology of finance and STS work to dismiss financial regulation as inherently ineffective, and encourages scholars to investigate its direct and indirect effects openly (see also Coombs, 2020).

It might be argued that our analysis is similarly rose-tinted: Our informants, at Tyler Capital and elsewhere, might have a self-serving interest in presenting their firms as being highly reliable and compliant. Given that high-frequency trading has received a substantial amount of bad press in recent years (e.g. Lewis, 2014), firms specializing in this form of algorithmic trading might be particularly inclined to give a positive impression of themselves. When conducting fieldwork at Tyler Capital, we deliberately and carefully crosschecked information across informants (across hierarchical levels and teams) and data sources (interviews, documents, observations). This led to and confirmed our conviction that the firm is genuinely committed to HRO principles, rather than treating this as mere window dressing. However, our analysis of Tyler Capital also demonstrates that even for a firm that is strongly committed to HRO principles and has gone further with this than other firms we encountered in our fieldwork, it still took staff almost 30 minutes to activate the kill switch in response to the high-speed cable network

outage. Compared to the response time of algorithmic trading systems, which are often measured in microseconds, this is not fast. Although the network outage was non-critical for the actual trading operations (a critical systems failure would likely have triggered a much faster response in most firms), this longer response time suggests that a full-scale HRO implementation in automated markets – and similar contexts, for that matter – may require comprehensive automated backing to avoid disasters (recall that it only took 45 minutes for Knight Capital to collapse). A similar observation is made by Scharre (2018) in his examination of automated warfare:

In fully autonomous systems, humans are present during the design and testing of a system and humans put the system into operation, but humans are not present during actual operations. They cannot intervene if something goes wrong. The *organization* that enables high reliability is not available – the machine is on its own, at least for some period of time. Safety under these conditions requires something more than high-reliability organizations. It requires high-reliability fully autonomous complex systems, and there is no precedent for such systems. (p. 178, original emphasis)

Algorithmic trading firms have an array of automated control systems in place, though not necessarily to the scale and extension recommended by Scharre. Although his call for a full integration of high reliability into the algorithmic systems – rather than having HRO principles implemented as an organizational wrapper around them – might be a yet-unattained goal, we believe that a proper organizational commitment to HRO principles may nonetheless contribute importantly to curtailing the escalation of both critical and non-critical incidents. However, such a commitment would obviously need to be aligned with the temporal operational scale of the technological systems (in this case, the trading algorithms), and from an HRO point of view, this would likely entail both automated and fast human response systems.

We are more concerned that, even if a market-wide implementation of HRO principles in trading firms would likely help curb incidents and crashes in (and their impact on) automated markets, such implementation may not fully address the extra-organizational and systemic aspects of these markets. Therefore, on the one hand, we agree with Bush et al. that ‘failures of high reliability in finance’ abound. Moreover:

The financial system is every bit as vital to our modern world as flight operations, space launches, firefighting, and industrial operations (all classic territory for high-reliability organizational, or HRO, studies), and it is necessary that constituent firms begin to behave in a high-reliability manner. (Bush et al., 2012: 169)

On the other hand, the discussion of a firm-level commitment to HRO principles needs to be combined with an appreciation of the NAT features of automated markets. This is important because, although our discussion of Tyler Capital suggests that this particular firm’s HRO implementations responded to its extra-organizational embeddings and connections, under certain conditions, intra-organizational commitments to high-reliability practices might lead to market instability paradoxically.

To demonstrate this point, we return to the Flash Crash and one of its technological issues, which concerns order book data feeds. Algorithmic trading firms generally rely

on two types of data feeds: (1) proprietary, direct data feeds from each individual exchange, and (2) public, consolidated data, consisting of information collated from all exchanges. The direct feeds are faster than the consolidated feed because time is not lost to the consolidation process. As MacKenzie (2021: 229) notes, ‘HFT firms do not rely on the relatively slow official datafeed for their trading, but they do use it as a data-integrity check on the faster direct feeds that they purchase from exchanges’. If the consolidated feed differs from the direct feeds, this is a reason for concern in trading firms, suggesting that their algorithms might be receiving incorrect data from the exchange(s) and therefore might be sending erroneous orders to the market.

As mentioned earlier, the NYSE experienced serious data feed delays on May 6, 2010 (CFTC–SEC, 2010: 77). Its own proprietary data feeds were only slightly delayed (on average about 0.008 seconds), but its technology for transmitting data to the public was slowed down, for several stocks, by more than 20 seconds on average. As a result, many trading firms activated what the CFTC–SEC refer to as ‘feed-driven integrity pauses’ (p. 36). Alarmed that their data-integrity checks showed inconsistencies between exchange feeds and the consolidated feed, firms automatically withdrew from markets, thereby exacerbating the liquidity crisis and driving markets further down. Although the CFTC–SEC (2010: 5, n. 10) argues that feed-driven integrity pauses did not play ‘a dominant role’ in the Flash Crash, researchers have challenged this and suggested that the data feed issue was indeed a critical part of the event (Aldrich et al., 2017). Building on this research, MacKenzie concludes that:

failed data-integrity checks ... might thus have led to the widespread shutting down of automated share-trading systems and to the disorderly trading that took place in their absence. If that is correct, it is a fascinating example of individually prudent, rule-bound behavior (checking data integrity *is* prudent) generating potentially damaging, unruly, collective turmoil. (MacKenzie, 2021: 229, original emphasis)

This is precisely the point we seek to generalize. Because of the tight coupling and complex interactions of automated markets, what constitutes precautionary high-reliability measures from the point of view of individual trading firms might paradoxically weaken the stability of markets, at least in some situations. During the Flash Crash, this manifested itself not only in feed-driven integrity pauses but also in the rapid price drops triggering trading firms’ internal risk control systems, leading them ‘to curtail, pause, and sometimes completely halt, their trading activities’ (CFTC–SEC, 2010: 36). One common risk control is stop-loss orders, designed to exit positions automatically when losses exceed a pre-set level. Like data integrity checks, stop-loss orders form part of the suite of measures one could demand by HRO-committed trading firms. However, while sensible from the point of view of an individual firm, in practice, stop-loss orders may generate a downward spiral when interacting with other stop-loss orders. When a price falls because of a trading firm’s stop-loss order, the lowered price can trigger other stop-loss orders, which in turn moves the prices further down, potentially triggering even more stop-loss orders, and so on. Since stop-loss orders stay hidden until the predefined trigger price is reached, firms have no information to anticipate such interactions between stop-loss orders. Furthermore, algorithmic trading systems trigger such orders at high speed, allowing little space for humans to intervene.

Our point is not to suggest that financial markets should revert to a state where market interactions take place on a human timescale, making room for humans to intervene if markets become unstable. The past couple of decades' technological and regulatory developments make such a reversal highly unlikely. Instead, we argue that debates about the stability of automated markets need to appreciate three central points. First, we need to appreciate that these markets are characterized by tight coupling and complex interactions, as illustrated by the Flash Crash and similar events. Second, although tight coupling and complex interactions make systems susceptible to normal accidents, HRO scholars have compellingly demonstrated that some organizations are capable of operating in technologically complex settings without causing large-scale accidents. These insights can be applied to algorithmic trading firms, but this requires, third, that HRO emphases be extended beyond intra-organizational dimensions and take seriously the ways in which algorithmic trading firms are connected to other organizations. In other words, given that automated markets are characterized by such tight coupling and complex interactions, HRO commitments cannot stand alone. Because of the coupling and interactions, there is an ongoing imminent risk that the implementation of HRO principles in individual trading firms may exacerbate rather than curb market instability – a point often ignored in the broader HRO literature.

Therefore, an appreciation of the need for both NAT and HRO insights is needed. This may translate into measures that strengthen HRO implementation within algorithmic trading firms (and trading venues), as well as more systemic measures that address the tight coupling and complex interactions of markets. The latter measures may come in many forms. For example, regulators may consider the level of leveraging used by trading firms. Since leveraging introduces tight coupling into the financial system (Guillén, 2015; Guillén and Suárez, 2010), trading on investor-provided rather than borrowed capital could potentially result in fewer market-wide crashes. However, given that the recent regulatory changes that have addressed the issue of overleverage have not alleviated market disruptions that originate from issues in technical systems (Gallagher, 2014), other measures that focus more directly on the organization of automated market participants might also be needed.

This could include requirements that designated market makers – trading firms that specialize in providing liquidity to markets and receive certain benefits from exchanges for that service – are obliged to keep trading, also when markets go against them. This was previously one of the key functions of the so-called specialists on the NYSE trading floors, who were obliged to step 'in as significant buyers in a falling market and as sellers in a rising market', even though this might take a heavy toll on their earnings (Mattli, 2019: 111). Although this type of system generated concerns about specialists mismanaging their duties and acting opportunistically, it has also been observed that, on May 6, 2010, those specialists who had survived years of market automation and were trading on the few NYSE floors still active, managed to curtail some of the turmoil during the Flash Crash (MacKenzie, 2015a). Again, we do not suggest a reversion to human trading, but it may be relevant to reconsider the rules that oblige algorithmic market makers to retain their presence in markets, including when they go through more turbulent phases. This is something that is missing from today's markets on a wide scale. For example, according to the standard agreement at the London Stock Exchange, derivatives market makers are

obliged to be present in the market for only 50% of a normal trading day (LSE, 2021). The same applies to the CME Amsterdam, but with the added note that companies are exempt from this obligation under ‘exceptional circumstances’ such as ‘a situation of extreme volatility’ (CME Amsterdam, 2021). In other words, market makers can withdraw from markets when, from a market stability point of view, they may be most needed, and where their withdrawal may exacerbate instability as under the Flash Crash.

We do not suggest that a stronger obligation around market makers’ presence would alleviate all the technological risks of automated markets – and we appreciate that requiring market makers to be more present in markets would have to be balanced with their need to activate the kill switch in certain situations such as the cable incident discussed earlier. However, market-maker presence is a tool among others, and one that precisely addresses the problem that, because markets are characterized by tight coupling and complex interactions, the implementation of HRO principles within trading firms may be insufficient, as an automated, HRO compliant, withdrawal of market makers – either in smaller flash crash episodes when volatility suddenly spikes, or in larger ones where market instability erupts quickly and spreads across markets – can aggravate the problems rather than ameliorating them. This touches upon a larger challenge: Even disregarding cases where the adoption of HRO principles might exacerbate disasters, HRO implementations, such as the one undertaken by Tyler Capital analyzed in this article, do not necessarily refute Perrow’s pessimism. Although an industry-wide commitment to an HRO-like risk-management culture would probably render normal accidents in markets less likely, the tight coupling and complex interactions of automated trading would still create fertile ground for normal accidents, even if they only occur infrequently.

## **Conclusions**

We have argued that present-day automated financial markets exhibit the characteristics emphasized by NAT scholars when pointing to accident-prone technological systems. Automated markets encompass a wide range of market participants (automated trading firms) whose interactions are complex and sometimes characterized by feedback loops. Further, trading algorithms are often designed to monitor and respond to changes in the exchanges’ electronic order books, with different exchanges and types of securities being tightly coupled. According to a NAT perspective, it is therefore not surprising to see market accidents in the form of flash crash events or firm-level collapses. Technological failures leading to large-scale accidents are expected in a system characterized by tight coupling and complex interactions. That said, we have also argued that the determinism of NAT need not be emulated. As HRO scholars have demonstrated, it is possible that the implementation of HRO principles will help curtail much of the technological risk otherwise characterizing automated markets.

Following this, our HRO analysis, based on our fieldwork, suggests that the traditional intra-firm emphasis in HRO scholarship does not sufficiently address the fact that trading firms are deeply interconnected with their environments. The central corollary of markets with tight coupling and complex interactions is that individual trading firms are not just exposed to internally generated incidents but also to externally produced ones, be it other trading firms’ erroneous orders (as in the Knight Capital case) or third-party

provider failures (as in our cable outage discussion). Accordingly, HRO principles need to be updated. Specifically, we suggest that high reliability in interconnected settings such as automated financial markets requires that firms systematically attend to – meaning, identify, respond to, and learn from – internal as well as external incidents that might affect their operations.

We also argue, however, that, although the implementation of HRO principles in individual organizations may help curtail technological failures and risk and should therefore be encouraged, pursuing HRO ideas locally might, paradoxically, weaken the overall stability of markets under certain conditions. The 2010 Flash Crash aptly illustrates this point insofar as precautionary measures made market participants withdraw from markets, which further exacerbated the liquidity crisis. This leaves us somewhere between the pessimism of NAT and the optimism of HRO scholarship, leading us to advocate a combined approach where HRO principles are joined by measures that address the NAT features of markets. We believe the adoption of such a dual analytical approach is helpful in both understanding automated markets and reducing the likelihood of future market failures and crashes.

### Acknowledgements

We thank Zachary David, Kjell Doksum, Neil Fligstein, Joan H Fujimura, Kristian Bondo Hansen, Robin Holt, Lise Justesen, Michael Pedersen, Justine G Pors, Karlene Roberts, and Daniel Souleles for their valuable comments and productive feedback on earlier versions of this article. We are also grateful to the *Social Studies of Science* Editor Sergio Sismondo and three anonymous reviewers for their insightful comments and helpful suggestions. We are immensely grateful to our informants, especially Chris Donnan, Mike Bushore, and the rest of the Tyler Capital community for granting us access to study their work. We would also like to thank Caitlin Welch for her research assistance.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 725706).

### ORCID iD

Bo Hee Min  <https://orcid.org/0000-0003-2800-7370>

### References

- Aldrich EM, Grundfest J and Laughlin G (2017) The flash crash: A new deconstruction. *SSRN*. Available at: <https://ssrn.com/abstract=2721922> (accessed 27 July 2021).
- Beunza D (2019) *Taking the Floor: Models, Morals, and Management in a Wall Street Trading Room*. Princeton, NJ: Princeton University Press.
- Beverungen A and Lange A-C (2018) Cognition in high-frequency trading: The costs of consciousness and the limits of automation. *Theory Culture & Society* 35(6): 75–95.
- Bigley GA and Roberts KH (2001) The incident command system: High-reliability organizing for complex and volatile task environments. *Academy of Management Journal* 44(6): 1281–1299.

- Borch C (2016) High-frequency trading, algorithmic finance and the flash crash: Reflections on eventalization. *Economy and Society* 45(3-4): 350–378.
- Borch C (2020) *Social Avalanche: Crowds, Cities and Financial Markets*. Cambridge: Cambridge University Press.
- Borch C and Lange A-C (2017) High-frequency trader subjectivity: Emotional attachment and discipline in an era of algorithms. *Socio-Economic Review* 15(2): 283–306.
- Bush NI, Martelli PF and Roberts KH (2012) Failures of high reliability in finance. In: Shrivastava P and Statler M (eds) *Learning From the Global Financial Crisis: Creatively, Reliably, and Sustainably*. Stanford, CA: Stanford University Press, 167–187.
- Callon M (1998) Introduction: The embeddedness of economic markets in economics. In: Callon M (ed.) *The Laws of the Markets*. Oxford: Blackwell Publishers, 1–57.
- Callon M (2008) Economic markets and the rise of interactive agencements: From prosthetic agencies to habilitated agencies. In: Pinch T and Swedberg R (eds) *Living in a Material World: Economic Sociology Meets Science and Technology Studies*. Cambridge, MA: The MIT Press, 29–56.
- Carroll JS (2018) Three lenses for understanding reliable, safe, and effective organizations: Strategic design, political, and cultural approaches. In: Ramanujam R and Roberts KH (eds) *Organizing for Reliability: A Guide for Research and Practice*. Stanford, CA: Stanford University Press, 37–60.
- CFTC-SEC [Commodity Futures Trading Commission and Securities and Exchange Commission] (2010) Findings regarding the market events of May 6, 2010: Report of the staffs of the CFTC and SEC to the joint advisory committee on emerging regulatory issues. Available at: <https://www.sec.gov/files/marketevents-report.pdf> (accessed 20 October 2010)
- Cliff D and Northrop L (2012) The global financial markets: An ultra-large-scale systems perspective. In: Calinescu R and Garlan D (eds) *Monterey Workshop 2012*. Switzerland: Springer Nature, 29–70.
- CME Amsterdam (2021) Market making agreement. Available at: <https://www.cmegroup.com/trading/market-tech-and-data-services/files/brokertec-market-maker-agreement-template.pdf> (accessed 6 September 2021).
- CME Group (2010) What happened on May 6? Available at: [https://www.cmegroup.com/international/files/May\\_6\\_2010\\_Incident\\_English.pdf](https://www.cmegroup.com/international/files/May_6_2010_Incident_English.pdf) (accessed 24 May 2010).
- Coombs N (2016) What is an algorithm? Financial regulation in the era of high-frequency trading. *Economy and Society* 45(2): 278–302.
- Coombs N (2020) Financial regulation. In: Borch C and Wosnitzer R (eds) *The Routledge Handbook of Critical Finance Studies*. New York: Routledge, 137–153.
- Curran D (2020) Connecting risk: Systemic risk from finance to the digital. *Economy and Society* 49(2): 239–264.
- Department of Treasury, Federal Reserve Board and NY Federal Reserve (2015) *Joint Staff Report: The U.S. Treasury Market on October 15, 2014*. Available at: <https://home.treasury.gov/system/files/276/joint-staff-report-the-us-treasury-market-on-10-15-2014.pdf> (accessed 27 July 2021).
- Eisenhardt KM and Graebner ME (2007) Theory building from cases: Opportunities and challenges. *Academy of Management Journal* 50(1): 25–32.
- European Parliament and Council of the European Union (2014) Directive 2014/65/EU of the European parliament and of the Council of 15 May 2014 on markets in financial instruments and amending directive 2002/92/EC and directive 2011/61/EU text with EEA relevance.
- Gallagher DM (2014) The philosophies of capital requirements. Available at: <https://www.sec.gov/news/speech/2014-spch011514dmg> (accessed 27 July 2021).

- Gephart RP, Van Maanen J and Oberlechner T (2009) Organizations and risk in late modernity. *Organization Studies* 30(2–3): 141–155.
- Golub A, Keane J and Poon S-H (2012) High frequency trading and mini flash crashes. SSRN: <https://ssrn.com/abstract=2182097> (accessed 6 September 2021).
- Guillén MF (2015) *The Architecture of Collapse: The Global System in the 21st Century*. Oxford: Oxford University Press.
- Guillén MF and Suárez SL (2010) The global crisis of 2007–2009: Markets, politics, and organizations. In: Lounsbury M and Hirsch PM (eds) *Markets on Trial: The Economic Sociology of the US Financial Crisis*. Bingley, UK: Emerald Group Publishing Limited, 257–280.
- Hansen KB (2020) The virtue of simplicity: On machine learning models in algorithmic trading. *Big Data & Society* 7(1): 1–14.
- Hansen KB (2021) Model talk: Calculative cultures in quantitative finance. *Science Technology & Human Values* 46(3): 600–627.
- Hansen KB and Borch C (2021) The absorption and multiplication of uncertainty in machine-learning-driven finance. *British Journal of Sociology*, 72(4): 1015–1029.
- Helbing D (2013) Globally networked risks and how to respond. *Nature* 497(7447): 51–59.
- Johnson N, Zhao G, Hunsader E, et al. (2012) Financial black swans driven by ultrafast machine ecology. arXiv:1202.1448.
- Johnson N, Zhao G, Hunsader E, et al. (2013) Abrupt rise of new machine ecology beyond human response time. *Scientific Reports* 3(1): 2627–7.
- Kirilenko A, Kyle AS, Samadi M, et al. (2017) The flash crash: High-frequency trading in an electronic market. *The Journal of Finance* 72(3): 967–998.
- Kirilenko AA and Lo AW (2013) Moore’s law versus Murphy’s law: Algorithmic trading and its discontents. *Journal of Economic Perspectives* 27(2): 51–72.
- Knorr Cetina KK and Bruegger U (2000) The market as an object of attachment: Exploring post-social relations in financial markets. *The Canadian Journal of Sociology/Cahiers canadiens de sociologie* 25(2): 141–168.
- Knorr Cetina KK and Bruegger U (2002) Global microstructures: The virtual societies of financial markets. *American Journal of Sociology* 107(4): 905–950.
- Lange A-C (2016) Organizational ignorance: An ethnographic study of high-frequency trading. *Economy and Society* 45(2): 230–250.
- Lange A-C, Lenglet M and Seyfert R (2019) On studying algorithms ethnographically: Making sense of objects of ignorance. *Organization* 26(4): 598–617.
- La Porte TR (1994) A Strawman speaks up: Comments on the limits of safety. *Journal of Contingencies and Crisis Management* 2(4): 207–211.
- LaPorte TR and Consolini PM (1991) Working in practice but not in theory: Theoretical challenges of ‘high-reliability organizations’. *Journal of Public Administration Research and Theory* 1(1): 19–48.
- Lenglet M (2011) Conflicting codes and codings: How algorithmic trading is reshaping financial regulation. *Theory Culture & Society* 28(6): 44–66.
- Leveson N, Cuthler-Gershenfeld J, Carroll J, et al. (2009) Systems approaches to safety: NASA and the space shuttle disasters. In: Starbuck W and Farjoun M (eds) *Organization at the Limit: Lessons From the Columbia Disaster*. Malden, MA: Blackwell, 269–288.
- Lewis MM (2014) *Flash Boys: A Wall Street Revolt*. New York, NY: WW Norton New York.
- LSE [London Stock Exchange] (2021) Market maker agreement. Available at: <https://www.lseg.com/sites/default/files/content/documents/CurveGlobal%20Markets%20-%20MiFID%20Market%20Maker%20Agreement%20%2824%20June%202019%29.pdf> (accessed 06 September, 2021).

- Maas MM (2018) Regulating for ‘normal AI accidents’ – Operational lessons for the responsible governance of AI deployment. In: *Proceedings of the 2018 AAAI/ACM Conference on Artificial Intelligence, Ethics and Society*, 2018.
- MacKenzie D (2006) *An Engine, Not a Camera: How Financial Models Shape Markets*. Cambridge, MA: The MIT Press.
- MacKenzie D (2015a) Dark markets. *London Review of Books* 37(11): 29–32.
- MacKenzie D (2015b) Mechanizing the Merc: The Chicago Mercantile Exchange and the rise of high-frequency trading. *Technology and Culture* 56(3): 646–675.
- MacKenzie D (2017a) A material political economy: Automated trading desk and price prediction in high-frequency trading. *Social Studies of Science* 47(2): 172–194.
- MacKenzie D (2017b) Capital’s geodesic. In: Wajcman J and Dodd N (eds) *The Sociology of Speed: Digital, Organizational, and Social Temporalities*. Oxford: Oxford University Press, 55–71.
- MacKenzie D (2018a) ‘Making’, ‘taking’ and the material political economy of algorithmic trading. *Economy and Society* 47(4): 501–523.
- MacKenzie D (2018b) Material signals: A historical sociology of high-frequency trading. *American Journal of Sociology* 123(6): 1635–1683.
- MacKenzie D (2019) How algorithms interact: Goffman’s ‘interaction order’ in automated trading. *Theory Culture & Society* 36(2): 39–59.
- MacKenzie D (2021) *Trading at the Speed of Light: How Ultrafast Algorithms Are Transforming Financial Markets*. Princeton, NJ: Princeton University Press.
- MacKenzie D and Spears T (2014a) ‘The formula that killed Wall Street’: The Gaussian copula and modelling practices in investment banking. *Social Studies of Science* 44(3): 393–417.
- MacKenzie D and Spears T (2014b) ‘A device for being able to book P&L’: The organizational embedding of the Gaussian copula. *Social Studies of Science* 44(3): 418–440.
- Mattli W (2019) *Darkness by Design*. Princeton, NJ: Princeton University Press.
- Menkveld AJ and Yueshen BZ (2019) The flash crash: A cautionary tale about highly fragmented markets. *Management Science* 65(10): 4470–4488.
- Palmer D and Maher M (2010) A normal accident analysis of the mortgage meltdown. In: Lounsbury M and Hirsch PM (eds) *Markets on Trial: The Economic Sociology of the US Financial Crisis*. Bingley, UK: Emerald Group Publishing Limited, 219–256.
- Pardo-Guerra JP (2019) *Automating Finance: Infrastructures, Engineers, and the Making of Electronic Markets*. Cambridge, MA: Cambridge University Press.
- Perrow C (1981) Normal accident at Three Mile Island. *Society* 18(5): 17–26.
- Perrow C (1984) *Normal Accidents: Living With High-Risk Technologies*. Princeton, NJ: Princeton University Press.
- Perrow C (1999) *Normal Accidents: Living With High-Risk Technologies*, 2nd edn. Princeton, NJ: Princeton University Press.
- Perrow C (2010) The meltdown was not an accident. In: Lounsbury M and Hirsch PM (eds) *Markets on Trial: The Economic Sociology of the US Financial Crisis*. Bingley, UK: Emerald Group Publishing Limited, 309–330.
- Pinch T and Swedberg R (eds) (2008) *Living in a Material World: Economic Sociology Meets Science and Technology Studies*. Cambridge, MA: The MIT Press.
- Ramanujam R (2003) The effects of discontinuous change on latent errors in organizations: The moderating role of risk. *Academy of Management Journal* 46(5): 608–617.
- Rijppma JA (1997) Complexity, tight-coupling and reliability: Connecting normal accidents theory and high reliability theory. *Journal of Contingencies and Crisis Management* 5(1): 15–23.
- Rijppma JA (2003) From deadlock to dead end: The normal accidents- high reliability debate revisited. *Journal of Contingencies and Crisis Management* 11(1): 37–45.

- Roberts KH (1989) New challenges in organizational research: High reliability organizations. *Industrial Crisis Quarterly* 3(2): 111–125.
- Roberts KH (1990) Some characteristics of one type of high reliability organization. *Organization Science* 1(2): 160–176.
- Roberts KH (1993) Cultural characteristics of reliability enhancing organizations. *Journal of Managerial Issues* 5(2): 165–181.
- Roberts KH (2018) Advancing organizational reliability. In: Ramanujam R and Roberts KH (eds) *Organizing for Reliability: A Guide for Research and Practice*. Stanford, CA: Stanford University Press, 3–16.
- Roberts KH and Libuser C (1993) From bhopal to banking: Organizational design can mitigate risk. *Organizational Dynamics* 21(4): 15–26.
- Roe E and Schulman P (2016) *Reliability and Risk: The Challenge of Managing Interconnected Infrastructures*. Stanford, CA: Stanford University Press.
- Sagan SD (1993) *Limits of Safety: Organizations, Accidents, and Nuclear Weapons*. Washington, DC: American Association for the Advancement of Science.
- Scharre P (2018) *Army of None: Autonomous Weapons and the Future of War*. New York: WW Norton & Company.
- Schneiberg M and Bartley T (2010) Regulating or redesigning finance? Market architectures, normal accidents, and dilemmas of regulatory reform. In: Lounsbury M and Hirsch PM (eds) *Markets on Trial: The Economic Sociology of the US Financial Crisis*. Bingley, UK: Emerald Group Publishing, 281–308.
- SEC [Securities and Exchange Commission] (2013) In the matter of Knight Capital Americas LLC. Available at: <https://www.sec.gov/whistleblower/award-claim/award-claim-2013-93> (accessed 6 September 2021).
- SEC [Securities and Exchange Commission] (2014) Regulation systems compliance and integrity; final rule. Available at: <https://www.govinfo.gov/content/pkg/FR-2014-12-05/pdf/2014-27767.pdf> (accessed 18 June 2016).
- Seyfert R (2016) Bugs, predations or manipulations? Incompatible epistemic regimes of high-frequency trading. *Economy and Society* 45(2): 251–277.
- Shrivastava S, Sonpar K and Pazzaglia F (2009) Normal accident theory versus high reliability theory: A resolution and call for an open systems view of accidents. *Human Relations* 62(9): 1357–1390.
- Sornette D and von der Beche S (2011) Crashes and high frequency trading. *Swiss finance Institute Research Paper No.* 11–63. SSRN 1976249. Available at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1976249](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1976249) (accessed 6 September 2021).
- Souleles D (2019) The distribution of ignorance on financial markets. *Economy and Society* 48(4): 510–531.
- Sutcliffe KM (2011) High reliability organizations (HROs). *Best Practice & Research Clinical Anaesthesiology* 25(2): 133–144.
- US House (2010) *The Stock Market Plunge: What Happened and What Is Next? Hearing Before the Subcommittee on Capital Markets, Insurance, and Government Sponsored Enterprises of the Committee on Financial Services*. Washington, DC: US Government Printing Office.
- US Senate (2010) *Examining the Causes and Lessons of the May 6th Market Plunge: Hearing Before the Subcommittee on Securities, Insurance, and Investment of the Committee On Banking, Housing, and Urban Affairs, United States Senate*. Washington, DC: US Government Printing Office.
- van der Veegt GS, Essens P, Wahlström M, et al. (2015) Managing risk and resilience. *Academy of Management Journal* 58(4): 971–980.

- Vaughan D (1996) *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. Chicago, IL: University of Chicago Press.
- Vaughan D (2005) Organizational rituals of risk and error. In: Hutter B and Power M (eds) *Organizational Encounters With Risk*. Cambridge, MA: Cambridge University Press, 33–66.
- Weick KE (1987) Organizational culture as a source of high reliability. *California Management Review* 29(2): 112–127.
- Weick KE and Sutcliffe KM (2001) *Managing the Unexpected: Assuring High Performance in an Age of Uncertainty*. San Francisco, CA: Jossey-Bass.
- Weick KE and Sutcliffe KM (2007) *Managing the Unexpected: Resilient Performance in an Age of Uncertainty*, 2nd edn. San Francisco, CA: John Wiley and Sons, Inc.
- Weick KE and Sutcliffe KM (2015) *Managing the Unexpected: Sustained Performance in a Complex World*, 3rd edn. San Francisco, CA: John Wiley & Sons.
- Williams TA, Gruber DA, Sutcliffe KM, et al. (2017) Organizational response to adversity: Fusing crisis management and resilience research streams. *The Academy of Management Annals* 11(2): 733–769.
- Williams TA and Shepherd DA (2016) Building resilience or providing sustenance: Different paths of emergent ventures in the aftermath of the Haiti earthquake. *Academy of Management Journal* 59(6): 2069–2102.
- World Economic Forum (2019) *Navigating uncharted waters: A roadmap to responsible innovation with AI in financial services*. Geneva, Switzerland: World Economic Forum.
- Young B (2012) Leadership and high-reliability organizations: Why banks fail. *The Journal of Operational Risk* 6(4): 67–87.

### Author biographies

**Bo Hee Min** is Assistant Professor of Economic Sociology at Copenhagen Business School. Her research examines how technology, algorithms, and artificial intelligence transform work, organizations, and markets.

**Christian Borch** is Professor of Economic Sociology and Social Theory at the Copenhagen Business School. He is the Principal Investigator of the European Research Council-funded research project ‘Algorithmic Finance’. His latest book is *Social Avalanche: Crowds, Cities and Financial Markets* (Cambridge University Press, 2020).