Beyond the Flash Crash: Systemic Risk, Reliability, and High Frequency Financial Markets

Andrew Kumiega
Adjunct Professor
Illinois Institute of Technology
565 W. Adams, Chicago, IL 60661
p. 630 484-3301 e. akumiega@iit.edu.

Greg Sterijevski
Illinois Institute of Technology
565 W. Adams, Chicago, IL 60661
p. 630 408-8022 e. gsteri1@hotmail.com

Ben Van Vliet, *Corresponding*
Assistant Professor
Stuart School of Business
Illinois Institute of Technology
565 W. Adams, Chicago, IL 60661
p. 312 906-6513 e. bvanvliet@stuart.iit.edu.

1/8/2016

Abstract

Extreme events in financial markets can arise from fundamental information, but they can also arise from latent hazards embedded in the market design. This is systemic risk and somebody bears this risk. These events add to risk and their probability and severity must be accounted for by market participants. This paper shows how this risk fits into the finance literature, and that from an engineering perspective this risk in markets has never been lower. The industry is evolving to mitigate this risk. This paper presents an overview of the complexity of the automated market network and describe how market participants interact through the exchange mechanism. It defines new terms and a new framework for understanding the risk of extreme market moves from a reliability and safety perspective.

Keywords: High frequency trading, reliability engineering, systemic risk.

Electronic copy available at: http://ssrn.com/abstract=2712768

Unlike the aftermath of other market snafus, such as the Knight Capital algorithm disruption in 2012, the botched Facebook IPO that same year, and the 2010 'flash crash', the three-and-a-half-*hour NYSE outage of July 8*[, 2015] *isn't being rehashed continuously with recriminations flying across CNBC*.

–Markets Media [2015]

Systemic risk (SR) in financial markets is the risk that the financial system (or a significant portion of it) will no longer perform its function (Poledna, et al. [2015]). This risk is often thought of as arising from a trigger failure (of a component firm in the system) that propagates or cascades through the system, causing other components to fail. In the listed markets for securities and derivatives, SR emerges from the synchronized behavior of (largely) algorithmic traders, the introduction of unintended order messages and transactions by out of control algorithms[1], or a cybersecurity breach. In listed markets, algorithmic agents interact in various machine-mediated ways and through different products and network paths (see Cozzo, et al. [2015]). Because algorithms are programmed to respond to the market activity of other algorithms, there is a potential for positive feedback looping, which could amplify changes and move the system away from its equilibrium state. The flash crash of May 6, 2010, is an example of this kind of synchronization. The Knight Capital[2] meltdown of August 1, 2012 is an example of SR from an out-of-control algorithm. While no known cybersecurity breaches have yet caused a market disruption, regulators are moving quickly to address the issue. In stark contrast, then, is the NYSE outage of July 8, 2015. This event shows that SR from component failures

[1] The interconnectedness of creditworthy agents is of less concern in listed markets as brokerage and central clearing guarantees many transactions.
[2] Knight Capital lost $440 million in firm value in roughly half an hour due to software problems in its trading systems (WSJ [2012]).

Electronic copy available at: http://ssrn.com/abstract=2712768

has been reduced over the years. Things are getting better. It had little, if any, effect on the ability of the market to perform its function.

Assessments of the ability of a financial system to perform its function are intimately related to concepts of quality, reliability, and safety, which are well-developed in the literature of industrial engineering. One of the central ideas in this literature is that hazards are embedded in the design of the system. Or, in other words, SRs are design flaws. One of the problems with analysis of SR in this arena is that the multiplex algorithmic network that is the listed markets consists not just of interconnected counterparties and the exposures among them, but also of agents that facilitate transactions between those counterparties—broker/dealers, clearing houses, exchanges, even independent software and hardware vendors. Failures of these agents' technologies can also trigger SR events, or what could be called market mishaps in the parlance of hazard analysis.

Extreme events are where reliability engineering and market risk intersect. An engineering perspective of risk in financial markets considers extreme events that arise either from the arrival of unforeseen information about fundamental values (say, of stocks), or from triggers of latent hazards embedded in the market design, including regulations and technology architectures. The flash crash is just such an event, the realization of a hazard risk. A latent hazard can be anything from a software bug in a trading algorithm to a macro-market system design that lacks redundant control measures. Some events propagate across the market network, as in the case of the flash crash, while others cause the firm to fail, but not the market itself, as in the case of Knight Capital.

Extreme events such as these add to the standard deviation of price moves, and their probability and severity must be accounted for (either explicitly or implicitly) by market participants. Someone bears the risk—either traders or society—through trades not consummated, higher implied volatilities, or wider spreads. Option traders must increase implied volatilities to account for unknown risks. Liquidity providers must charge a higher premium for their adverse selection risk. Yet, no methodology exists to assess the costs of (what is essentially) mishap risk in the markets.

In this paper, we attempt to describe the multiplex algorithmic network and develop the concepts of quality, reliability, and safety[3] for it. We present data that shows a "quality and reliability revolution" in financial markets is already well underway, and has been for at least five years. The industry appears to already be well beyond Six Sigma. Rather than policing quality, society would be better served if regulators moved on to focus on increasing the reliability and safety of the complex market network.

Architecture of Liquidity Supply and Demand

In this section, we review the architecture of automated liquidity supply chain, from those that supply liquidity to those that demand it. Virtually the entire liquidity supply chain is automated, so it's no exaggeration to say that essentially all of finance is now automated, and "automation is an engineering discipline (Schuler [2006])." Liquidity demanders are those market participants (or actors) who are willing to pay a fee (i.e. the bid-ask spread) for transactional immediacy. Liquidity suppliers are those who seek to earn the fee but incur the risk of waiting to transact. By and large, high frequency traders (HFT) using capable trading strategies[4] are the suppliers of liquidity.

[3] Quality engineering (or quality assurance) attempts to prevent defects in processes that deliver products or services to customers. It is primarily concerned with the development and monitoring of policies and procedures that ensure that the product or service meet customer requirements. Reliability engineering, on the other hand, attempts to improve the ability of a system to "adequately perform its specified purpose for a specified period of time under specified environmental conditions" (Leemis [1995]).

[4] A capable trading strategy is one that generates returns that consistently exceed its costs with some level of certainty (see Kumiega, et al. [2014]).

LFT and HFT actors use similar technological components, but in different configurations and with different levels of computational power and sophistication. Because reliability analysis in part consists of the assessing the "reliability importance" of components (Wang, et al. [2004]), we review agents' configurations' with respect to five components.

Strategy Engine Component ①: This component generates limit or market order requests. In the LFT case, this could be an individual entering an order through a retail website, or a large institutional money manager running an automated order entry algorithm (usually called an execution algorithm). In the case of the liquidity provider, this is usually a co-located, HFT strategy.

Risk Gate Component ②: This component is a computer server that performs validation and risk management checks of order requests. For HFT systems, this component is often embedded in the order management system (OMS). The incentive has been to omit this component because of the additional latency (or slowness) that it may add to the order entry process. However, as the industry has matured, almost all firms now incorporate this component into their co-located technology.

Broker/Clearer Component ③: This component is a server hosted by the broker/dealer or clearing member firm that guarantees the trades made by their customers. Such firms may be liable for the losses of their customers and so they monitor customer activity. Should risk limits be exceeded, the broker or clearing member can shut down the customer's trading in real-time. High frequency trading firms often seek to bypass this component with direct market access (DMA) due to the additional latency incurred.

Match Engine Component ④: This component is a server that resides at the execution venue (usually, an exchange) and operates the limit order book for the particular stock or derivatives contract. It includes the trade matching engine, which matches buyers and sellers orders together to make trades according to an algorithm, usually either first-in-first-out (i.e. FIFO, and according to price and time priority) or pro-rata (i.e. price and size priority).

Market Data Feed Component ⑤: This component is a server at the exchange that broadcasts data about all changes to the limit order book as a result of incoming order requests. Market data contains information about trade executions, and additions to and cancellations from the limit order book. This data is transmitted in real-time to all market participants that subscribe to the exchange UDP multicast. This data feed is the output process of the execution venue. Mishaps that occur in the market are transmitted to other participants by this data feed.

Exhibit 1 shows (at a high level) the architecture of an HFT system that supplies liquidity using the components just described. Exhibit 2 shows (at an equally high level) the architecture of LFT orders using the same components. As mentioned, the two architectures differ primarily in that HFT liquidity providers typically have[5] direct market access (or DMA).

In Exhibit 1, an HFT system that generates trading decisions (Component ①) exists in software on a server that is co-located at the exchange. The system receives price data from the exchange and runs its strategy decision logic. The home office of the trading firm receives data from this trading system about its performance and enables setting of parameters (e.g. on/off switches). This provides for external control of the system, but not pre-trade risk checking. Should an HFT system go haywire, these external controls can only be applied after the fact. As

[5] As discussed firm with DMA is connected directly to the matching engine component, without first going through a broker/clearing member component.

mentioned, most firms now have real-time risk checks embedded in Component ②. Trading decisions pass to the order management system (OMS), which performs these checks and controls on critical characteristics, such as the number of messages to the exchange, number of transactions, and network latency (see Bilson, et al. [2010]). The OMS routes limit and market order requests to the exchange order server (Component ④). (This is where the NYSE failure occurred.) When the exchange matching engine pairs (say) an HFT's limit buy order with a market sell from another participant (or vice versa), the order server sends a fill confirmation message back to the OMS to that effect. As discussed, messages regarding the activity in the limit order book are also broadcast to all subscribers from its price server (Component ⑤).

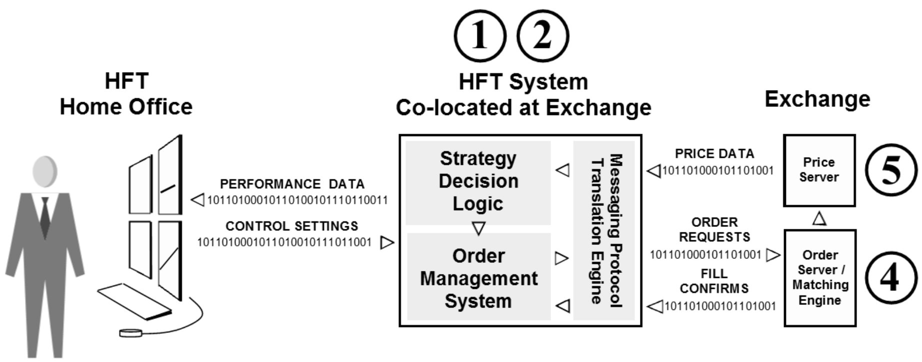

Exhibit 1: HFT Liquidity Suppliers Architecture

In Exhibit 2, LFTs generate their orders from the trader's desk either by pointing and clicking or through an algorithmic execution strategy (Component ①). These orders then (usually) pass through an internal server (Component ②), which performs risk checks, to a broker or clearing member (Component ③), which performs additional risk checks. Then, the broker or clearing member, who has direct market access, routes the order to the exchange (Components ④ and ⑤)[6].

[6] Some customer orders are crossed by the broker or routed to off-exchange execution venues.

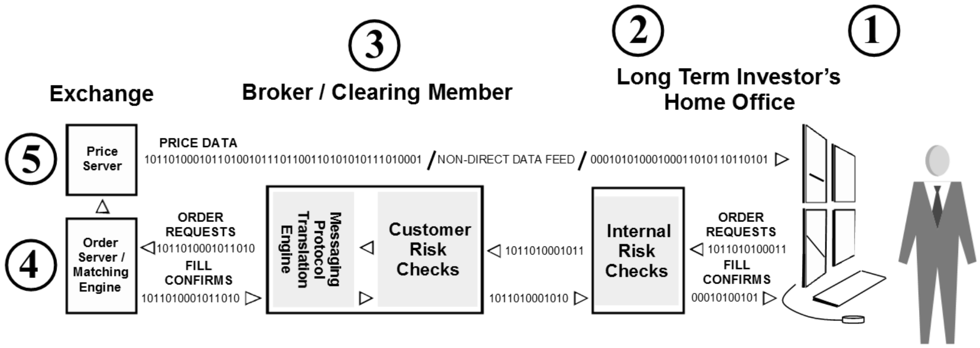

Exhibit 2: LFT Liquidity Demanders Architecture

Exhibits 1 and 2 depict the process flow of messages to and from HFTs and LFTs through the exchange. The flows of messages from all participants coalesce in the exchange limit order book and discover the equilibrium price. The process flow also determines the steady state for messaging (i.e. limit order book adds and cancels) within and between (or across) exchanges. Equilibrium prices and messaging steady states are spontaneous orders arising from self-organizing market networks[7]. From this perspective, the exchange limit order books are competitive process flow arenas, where the flow of messages from one participant causes other participants to react in a variety of ways according to their individual strategies. Exhibit 3 is an attempt to combine multiple Exhibits 1 and 2 to depict the complexity of this process flow arena. Within this arena, there are millions of trades and hundreds of millions of messages per day back and forth between the agents and their components.

In Exhibit 3, each HFT firm has a system co-located at more than one exchange. While multiple co-locations of order management systems are viewed as constituting a potentially predatory practice from a regulatory viewpoint since HFT firms route orders to multiple exchanges or engage in inter-exchange arbitrage. It is clear from the NYSE disruption that the

[7] This is why we ought to use agent-based models and simulation to understand the reliability and safety of financial markets.

ability to immediately route orders to other exchanges kept the market running with few to no issues. Therefore, what is seen as predatory by regulators, can be viewed as higher reliability from an engineering perspective. Even with only three HFT firms providing liquidity, three exchanges, two brokers/clearing members, and five LFTs demanding liquidity, the complexity of the market network is apparent. Of course, the real world markets have many more trading firms, dozens and dozens of exchanges and other execution venues, many more brokers and clearing members, and millions more long term investors. If we add in the millions of financial instruments listed on the exchanges, it is no exaggeration to say that the real markets are infinitely more complex than Exhibit 3.

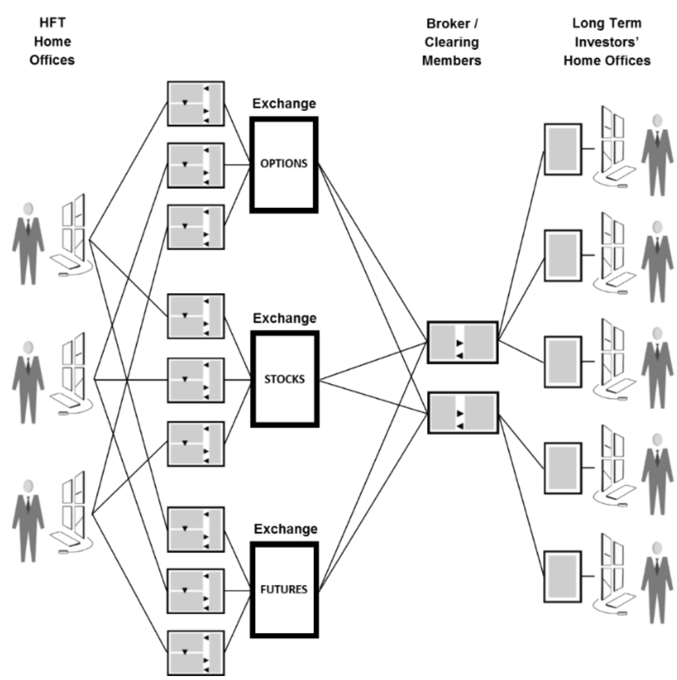

Exhibit 3: The Complex Automated Financial Market Network

Specifically, six additional factors increase the complexity of the network shown in Exhibit 3:

1. Most HFT firms are connected to multiple equity execution venues, multiple options exchanges, and multiple futures exchanges and engage in arbitrage between them.

2. Most brokers are connected to multiple equity execution venues (including dark pools, not just stock exchanges) and may route their customer orders to multiple venues in order to provide best execution.
3. This multiplicity applies across all exchanges, which further complicates the network design for each firm. The five components can be configured in virtually an infinite number of ways, which makes the effects of a market design change (such as a new regulation) at best uncertain since there is no standard configuration.
4. HFT firms, as well as brokers and clearing firms, do not disclose their network configurations or their trading algorithms.
5. Systems are reconfigured everyday with new hardware technologies—servers, cables, microwaves, lasers—and new and upgraded software.
6. Messages on one exchange can have a feedback loop to other exchanges. Changes in stock quotes can move futures quotes and vice versa.
7. If we overlay the regulatory regime on top of the physical, technological network structure, the complexity increases tremendously. Regulations (which are part of the market design and may require certain interactions between agents, order types, trading limits, and restrictions or obligations on trading algorithms) embed latent hazards into the market network.

In the complex financial network, the market output—the limit order book price feed (Component ⑤)—has become highly unpredictable and subject to feedback effects. The complexity leads to pricing and messaging dynamics that may (or may not) indicate market mishaps, or for that matter nefarious activity.

Definitions for Financial Markets

In this section we define some safety and reliability terms for the financial markets based upon those found in Ericson [2012].

- A *market hazard* (MH), any real or potential condition embedded in the components of the market or the market design that can cause financial loss or market disruption.
- A *market disruption* or *market error* is a failure of Component ④ and/or Component ⑤. It is an uncontrolled mishap.
- A *market mishap* (MM) is an unplanned event or series of events resulting in financial loss or market disruption. MMs arise from market hazards.
- *Systemic risk* or *market mishap risk* (SR or MMR) is an expression of the probability of a mishap occurring and its potential severity and the probability of controlling it.
- *Market system safety* (MSS) is the application of engineering and management principles, criteria, and techniques to achieve acceptable SR, within the constraints of operational effectiveness and suitability, time, and cost.
- While the term market quality has various meanings, we define *market quality* (MQ) is a snapshot measurement of the number of market disruptions or errors per quantity of market messages or transactions per unit of time. It is a measure of the normal variation in the market process. We note that design flaws that introduce latent hazards are always the root causes of poor market quality. These errors may be small and represent extremely low mishap risk. A large and uncontrolled error can cause a market mishap that contains a large probability of propagating across the market.
- *Market reliability* (MR) describes the ability of the complex market system to function over time. Reliability engineering realizes mishaps due to embedded hazards inevitably exist and aims to design processes to run despite them. A system which experiences

mishaps can be made to be highly reliable by controlling outcomes from latent hazards that cause mishaps by placing controls to insure the MMR is small. The level of error reduction points to redundant control systems that are reducing uncontrolled mishaps, as we will show.

- *Market safety* (MS) is the ability of the market to avoid financial loss or market disruption in the event of a mishap. An uncontrolled mishap can be a safety event, and they are almost always the root cause for large market crashes. Market events, including fundamentally driven extreme events, should not cause the market system to fail in a random manor. In a reliable market, mishaps are controlled to a fail-safe condition.

Systemic Risk and Volatility

Price volatility in financial markets arises from uncertainty about value. Uncertainty may increase when new information arrives and persist until all new expectations are incorporated. This is the traditional definition of risk as the standard deviation σ of returns. But, this conceptualization does not represent total process variation described in ISO/IEC Guide 73. SR must also increase uncertainty. As we have argued, in *algorithmic* financial markets, it must be the case as in equation (1) that total implied volatility σ_T is a mixture of uncertainty about price σ_P and SR σ_{SR}, which we define in Section 8. Furthermore, it certainly should be the case that there is an interaction, or correlation term, between them. It seems intuitive that more volatile markets are more prone to MMs.

$$\sigma_T = \sqrt{\sigma_P^2 + \sigma_{SR}^2 + 2 \cdot \sigma_P \sigma_{SR} \rho_{P,SR}} \quad (1)$$

The market design should mitigate this second form of risk. Even in the face of extreme market news, the market should remain reliable. If the market design evolves in a direction that makes any form of risk not hedge-able, trading firms will curtail their activities. This in turn will exacerbate volatility by creating liquidity gaps. By creating a more fragile (i.e. less reliable or safe) system, price risk becomes more expensive to manage and, therefore, will increase volatility.

Market mishaps disrupt the updating of the exchange order book (Component ④) and/or the dissemination of real-time market data (Component ⑤). Either of these can lead to increased variance of price changes. In engineering, quality and reliability are intertwined from the initial design of the process. As the process becomes high and higher quality through better design, eventually the Kaizan process shifts from percentage of, for example, machined parts within specification to, for example, zero defect days. The same should be true in the markets. Market stability is should not be the goal, but rather market reliability.

Extreme market moves often fall under the definition of unexpected tail events, and various mechanisms exists to account for their occurrence, such as mixture models. The second process is the unknown mishap risk. These unknown risks include things like a software bug in some firm's Component ① that causes unintended orders to reach the exchange limit order book (Component ④), network interruptions that prevent automated order management Component ②, the mishap of a competitor's system, unexpected interactions in the process flow of messages in the limit order book Component ④ (as in the case of the flash crash), and limit up limit down halts that halt activity in the limit order book. An uncontrolled mishap in one Component ①, the trigger mishap, can have a cascading effect through the process flow arena of Exhibit 3. These events may cause unexpected losses far outside the historical return distribution. Such mishaps, however, are rare despite what the media would like us to believe. If he output process is the broadcast from Component ⑤, then one back-of-the-napkin way to calculate the mishap rate in that process is to take the number of messages sent in error divided by the total number of messages.

For example, the CME's Globex system (Component ④) alone now handles over 300 million incoming messages per trading day from HFT and LFT participants which are broadcast through Component ⑤. If we assume that this number is representative of each large exchange, of which (let's just say) there are five in the world, the total number of messages per year might be something in the neighborhood of 3.8 trillion. From a cursory scan of the news, it appears that from 8/1/2011 to 7/31/2012 there were 31 possible mishaps worldwide. (This is a time period the media seemed to think was of low quality.) If a single error contains 10,000 bad messages, then the total number of errors is about 300,000 messages per year. Thus, the calculation of the error rate is 300,000 / 3.8 trillion = 8.20×10^{-7}. The automated financial network ought rightly to be called a Six Sigma industry[8]. Its quality exceeds the Six Sigma threshold of 3.40×10^{-6} by a wide margin. But, that doesn't mean quality, reliability, and safety cannot be improved, if only there were standard ways to measure these characteristics.

Systemic Risk and Bid-Ask Spreads

The bid-ask spread in financial instruments represents the transaction cost to liquidity demanders paid to liquidity providers. The literature assumes liquidity demanders pay half the bid-ask spread (or effective spread) as the fee to liquidity providers both when entering into and exiting from longer term positions (see Bessembinder [2003]; Hendershott and Riordan [2013]). The effective spread *epread* is most often decomposed into an estimate of the revenue to the liquidity providers, or realized spread *rspread*, plus an adverse selection (AS) premium, which is

[8] In quality engineering, a six sigma process is one in which 99.99966% of the outputs of a process are statistically expected to be free of defects. This works out to 3.4 defects per million. The calculation for a six sigma process is 3.4 / 1,000,000 = 3.40×10^{-6}.

essentially an insurance premium paid to liquidity providers for the risk they bear when passively accepting positions put to them by informed traders. However, as we have argued, it must also be the case that liquidity providers bear the risk for market mishaps, or SR, as well. Thus, the total adverse selection premium should be amended to include a premium for uncertainty about SR premium. Thus, we can decompose *espread* a step further into equation (2).

$$espread = rspread + AS\_premium + SR\_premium \qquad (2)$$

As with price volatility, the market design should mitigate this second form of risk. If the market design increases the probability that passively taken positions will not be hedge-able, then liquidity providers will respond by increasing the *SR_premium* charged to liquidity takers.

Quality in Financial Markets

With respect to (1) and (2), we note that mishap risk has been quickly declining. Redundancies have made some market failures, such as the NYSE outage of July 8, almost inconsequential. The financial markets are already a high quality industry. It appears to be beyond Six Sigma even in the time period commonly thought of as being poor quality. If this is true, then we would expect to see a decreasing number of mishaps, zero defect months, and increasing reliability. In this section, we present some empirical evidence generously provided by the Financial Information Forum (FIF). While the data presented is limited, from an industrial viewpoint, it is all the data that is in the public domain. This is due to the fact that both the HFT firms and the exchanges are highly proprietary and secretive with any data (see for example CNBC [2014]). The data shows that the financial industry has evolved from high quality to high reliability over the past 5 to 7 years. The industry has gone up the same curve that many other industries have, only faster.

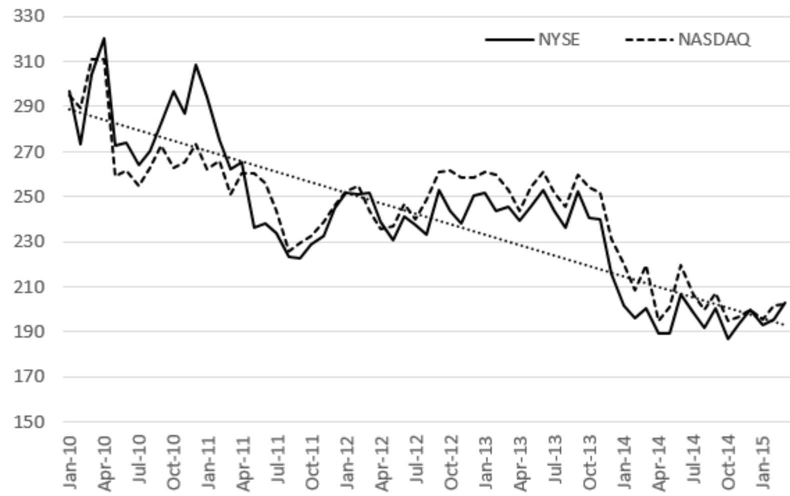

Exhibit 4: U.S. Equities Market Average Trade Size

In Exhibit 4, we see that the average trade size has decreased by approximately 50% since 2010. By and large, traders seem to prefer smaller orders that carry lower risk. But, we can also see three distinct time periods in Exhibit 4. The first is the period of declining trade sizes through roughly July, 2011, followed by a period of flat average trade size through October, 2013. After a sharp decrease, the average trade size has leveled off since January, 2014.

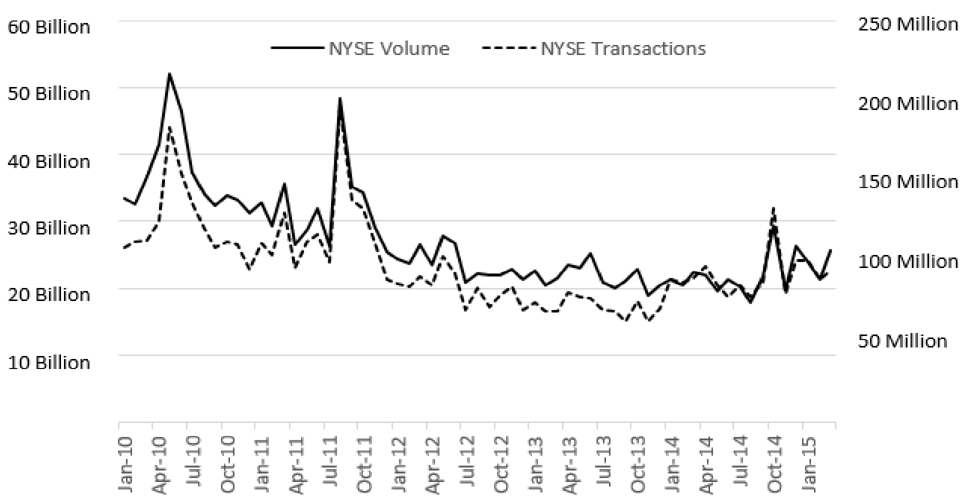

Exhibit 5: NYSE Total Volume and Transactions

Concurrently, we can see in Exhibit 5 that the transaction volume and total volume declined through the middle of 2012. In particular, though, transaction volume has been increasing since January, 2014, which corresponds to beginning of the lower average trade size regime in Exhibit 4. However, over the same time period, 2010-2014, while the number of transactions has been relatively flat since 2010 (though increasing more recently) and the average trade size decreasing, the number of self-declared (Component ②) mishaps has been dramatically decreasing, as can be seen in Exhibit 6. Until the NYSE event, the concept of self-help was understood by only the few firms with DMA. After the NYSE incident, the world should recognize the importance of this automatic messaging system that allows participants to route around a problem. In the data, we can see that any decrease in mishap rate in the market is not due to a decreasing amount of trading. If anything, we see the exact reverse. The recent upturn in number of transactions is increasing, but the number of mishaps is decreasing.

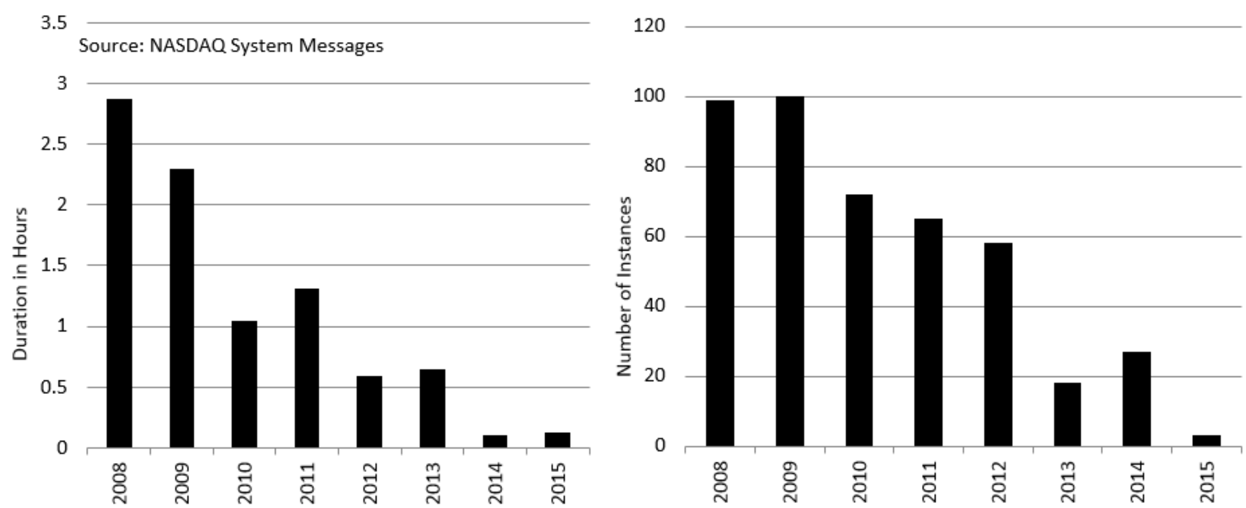

Exhibit 6: U.S. Equities Market NASDAQ Self-Help Declarations, Annual

One of the goals of reliability and safety engineering is zero-(Component ②) mishap months, and we can see in Exhibit 7 that the industry is now achieving this goal. This is another sign of high quality per year and increasing reliability over the years. What we are seeing is the result of quality engineering programs in the trading industry, with less frequent mishaps that cause market disruption and less severe consequences (at least in terms of duration) when mishaps do happen.

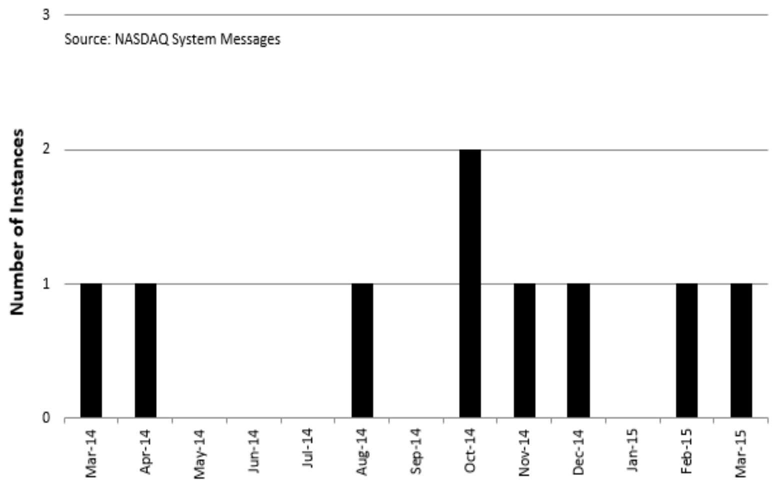

Exhibit 7: U.S. Equities Market NASDAQ Self-Help Declarations: Number of Instances

Quality engineering allows for consistent audits and an appropriate governance structure in order to mitigate the occurrence and severity of component mishaps, and quality engineering practices are now being applied within firms and exchanges. Yet, the media and the regulators seem to think the financial market network is low quality, and that the industry needs policing. However, during the NYSE event, the media was discussing how this was a non-event, but not recognizing the fact it was only so due to the reliability of the new systems. The quality-related stories in the press and regulatory actions are targeting a process that is already in a high quality steady state.

Reliability in Financial Markets

While largely synonymous, reliability is really the better term to use than quality. The reliability of a system is the probability that it will be able to perform its intended function over time. It may also be described in terms of a probability of a mishap, the frequency of mishaps, as in Exhibit 4, or in terms of availability. Using the data in Section 6, we can define what should correctly be called measures of market reliability. If we divide the duration of Component ② self-help instances by month times the duration of months *m*, we can arrive at a measure of market reliability *MR* as in equation (3).

$$MR_t = (m - SH_d) / m \tag{3}$$

Or, if we divide the number of self-help instances by the total number of transactions, we get another measure of market reliability level as in equation (4).

$$MR_n = (transactions - SH_n) / transactions \tag{4}$$

Say for example, in the old days, a trader in a trading pit engaged in 1,000 trades over the year, and that ten of those trades were out-trades[9]. For that trader, his or her reliability would be (1,000 – 10)/1,000 = 0.99, or a rate of 99%. The mishap rate is 1%, or expressed in orders of magnitude 1 per 100, or 1 per 10^2. Incidentally, up until the mid-2000's, many exchanges had special sessions, where out-trade clerks traded the mishaps. These sessions no longer exist in the electronic world. In Exhibit 8, point A represents this pit trader defect rate, while at point B the electronic market's error rate is something in the neighborhood of 1 per 10^8.

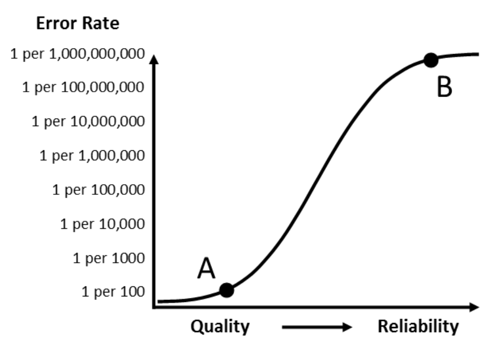

Exhibit 8: Quality, Reliability and the Error Rate of a Process

[9] An out trade occurs when there is a disagreement between the buyer and the seller in the trade. Because out trades were so common, the futures exchanges used to have an out trade session after the close of trading every day.

With respect to regulator's efforts to improve markets, rules aimed at increasing market quality can increase the reliability of market components. However, at the level of quality the markets are already at, such increases will likely not be significant. To increase reliability, regulators need to move beyond component-based compliance rules for firms, and focus on network reliability rules, redundancies, and consider extreme events that may arise from latent hazards in the market design. We add to this that the outcomes of new quality-driven regulations are not and cannot be fully understood. Regulators writes rules. They apply rules. And, then they hope the rules result in a desirable steady state, one that improves reliability. Without a testing framework, they cannot verify whether or not a permanent correction fixes the root cause. They are not addressing the complexity of market problems using the well-known tools of reliability engineering.

Regulators are focusing on containments for HFT through the Control Settings and the throttling of Order Requests in Exhibit 1. They are also focusing on the quality of Component ① and ② software code that launches individual order requests. They also aim to enhance fairness of execution prices across exchanges in Exhibit 3 through Reg. NMS[10]. Proposed regulations constitute market design changes and their impacts ought to be verified. Such design changes should not introduce other hazards. Unfortunately, this approach is not considered. There is a non-zero probability given the complexity in Exhibit 3 that a regulatory change will result in substantial decrease of market reliability.

And, this is a problem in academic finance as well. To address unknown risks in the market, we cannot use stochastic calculus or the efficient markets hypothesis. There are no elegant empirical justifications. The mishap distributions are unknown and unstable.

[10] The U.S. Securities and Exchange Commission's "Reg. NMS" is intended to assure that investors receive the best prices by encouraging competition between execution venues.

Addressing reliability and safety in markets requires a new perspective. From the reliability perspective, the markets have already achieved steady-state distributions with respect to liquidity, bid-ask spreads, and execution times. Before changes are introduced into the market design, we must understand how it will change the steady state—lower or higher—and how long it will take to get to the steady state. Only agent-based simulations will enable investigations of how changes will affect the market's steady states. Such simulations based on the market framework discussed will (likely we believe) show that some regulatory changes will have outcomes that do not promote the societal interest in greater market reliability. Such models will need to account for multiple dependencies across Exhibit 3, including the various types of Component ④ matching engines, in order to answer what-if questions. This is not easily done, but it is the only way for markets to evolve in a way that benefits individual investors and society.

Safety in Financial Markets

In the market network, zero component ② mishap risk is impossible[11]. For this reason, safety integrity levels (*SIL*) (see Charlwood, et al. [2004]) should be required of all components. *SIL* frameworks[12] typically rely on subjective judgements of unmeasurable parameters and probabilities, and can assign these values[13]:

1. The probability P_E of an unexpected participant or market mishap event E occurring.
2. The magnitude or severity S_E of the possible loss due to the event E.
3. The probability of controlling the loss given event E has occurred is given as $P_{c|E}$.

[11] This assumption is part of IEC 61508.
[12] For much more information, see IEC 61508.
[13] This perspective on risk is not unfamiliar in finance. Jarrow, et al. [1997] use a similar approach to modeling credit risk.

As in most *SIL* implementations, estimated scores can be assigned, say, from 1 to 4. For the severity S_E, a score of 4 may represent the highest, or intolerable, level of severity (like a flash crash), whereas a score of 1 may represent some trivial risk. A P_E score usually represents some range of probability of mishap, say 1 in 10^4 to 10^5. A score of 4 might mean highly likely; a score of 1 highly unlikely. A $P_{c|E}$ score of 4 might mean uncontrollable, while a score of 1 means almost completely controllable. Using these values, market participants could calculate *SIL*s as in equation (5)[14].

$$SIL_E = P_E \cdot S_E \cdot P_{c|E} = \sigma_{SR} \tag{5}$$

Let's assume an example scenario where the OMS of some HFT (Component ②) sends thousands of unintended sell orders into the market within a very short amount of time. Presumably, this is an intolerable risk for the firm and external market participants as well. The probability of this event occurring P_E on any given trading day may be fairly small. The severity of the loss S_E, however, could be very large. If the probability of control $P_{c|E}$ is very high, then the safety integrity level for that event SIL_E could meet some acceptable threshold. Considering all the possible independent mishap events for any given component in Exhibit 3, with all their corresponding severities and probabilities of occurrence and control, we can arrive at an *SIL* for each component and the entire network.

Alternatively, agent based simulations of the market could be performed. Given one simulation with SR and one without it, the difference in variance between to the two could be used as an estimate the effect of SR on σ_T. This approach is commonly used in industrial

[14] Equation (1) is adapted from the Automotive Safety Integrity Level (ASIL) risk classification scheme defined in ISO 26262 Functional Safety for Road Vehicles.

engineering to model, for example, supply chain disruptions (see Tomlin, 2006, and Kleijnen, 2005).

SIL models imply two ways to increase market safety. First, participants can build better quality systems with lower probabilities of mishap P_E. Second, they can build better real-time control systems that increase $P_{c|E}$. To increase safety, all firms whose systems pose such a risk in the event of their mishap ought to engage in both practices. We note that regulations now exist, or are proposed, that define rules for design and development (i.e. decreasing P_E), operation and control (i.e. increasing $P_{c|E}$) of individual components of automated financial markets that pose a systemic risk in the event of their mishap. This is also the perspective of industry-defined best practices, and the proposed quality management system standard for automated trading, ANSI X9 D13, more commonly known as AT 9000 (Van Vliet, et al. [2014])[15].

What is of concern, however, is also the reliability and safety of the overall market, not just the safety of the individual components. Market mishaps can arise from unexpected interactions between otherwise safe components. Increasing the reliability of the market requires lowering the probability of market mishaps $P_{E,m}$. This ought to be the focus of future research. This is important because removing all mishaps from the market is impossible, though regulators to often seem to think this is only goal.

Conclusion

The purpose of a financial market is to process transactions in order to facilitate price discovery, a societally beneficial function, which leads to efficient capital formation and risk transfer. Even an isolated single mishap in the market, such as the flash crash or the Knight

[15] ANSI (American National Standards Institute) is the U.S. component of ISO (International Organization for Standardization). X9 is that component of ANSI tasked with overseeing standards in the financial industry. The AT in AT 9000 stands for automated trading.

Capital meltdown, creates the perception of low reliability, which may cause anxiety among individual investors. This may increase implied volatilities and bid-ask spreads. As we have shown, based on the billions of messages sent every trading day, the financial markets are maybe the most high quality industry on earth. We show in this paper that the markets have moved from high quality to a reliability level that is on par with the utility industries. The NYSE event confirms our hypothesis that the markets have become reliable through the use of parallel systems, similar to power grids, which was not the case only a few years ago. Even high quality, complex systems can have catastrophic mishaps. In order to mitigate mishap risk, the industry needs to focus on how to increase reliability and safety without inadvertently jeopardizing the current level of reliability. Reliability and safety engineering practices must now be applied across the entire liquidity supply chain. If one aims to increase market reliability and safety, or assess the impact of changes in market design on these characteristics, the only mechanism available would be full-scale simulation (Black and Mejabi [1995]).

To improve the reliability and safety of the automated financial markets, new methods and tools need to be developed. These include large-scale methods for simulating dynamic financial systems at various scales. The solution to creating fail safe system is not policing but using simulations to test the various network, queuing and system safeties to create regulations that increase the reliability of the overall market in a reliable manner. This is the simulation-based model approach the Federal Energy Regulatory Commission (FERC) uses to regulate energy trading and the power grid. Regulating the network and the matching engines in financial markets is similar to regulating power lines, production, and usage.

References

Bessembinder H., “Issues in assessing trade execution costs.” Journal of Financial Markets. 6, (2003, pp. 233–257.

Bilson, J. F., Kumiega, A., Van Vliet, B. “Trading model uncertainty and statistical process control.” Journal of Trading. 5, (2010), pp. 39-50.

Black, J.J., Mejabi, O.O. “Simulation of complex manufacturing equipment reliability using object oriented methods.” Reliability Engineering and System Safety. 48, (1995), pp. 11-18.

Charlwood, M., Turner, S., Worsell, N. “A methodology for the assignment of safety integrity levels (SILs) to safety-related control functions implemented by safety-related electrical, electronic and programmable electronic control systems of machines.” UK Health and Safety Executive Research Report 216. (2004).

CNBC. “Citadel CEO: HFT firms need to be secretive.” (2014). Accessed on July 30, 2015 at: www.cnbc.com/2014/07/17/citadel-ceo-hft-firms-need-to-be-secretive.html

Cozzo, E., de Arruda, G.F., Rodrigues, F.A. and Moreno, Y. “Multilayer networks: metrics and spectral properties.” In Interconnected Networks, edited by F. Schweitzer and A. Garas. (2015). Springer.

Ericson, C. Hazard Analysis Primer. (2012). CreateSpace Independent Publishing.

Hendershott, T., Riordan, R. “Algorithmic trading and the market for liquidity.” Journal of Financial & Quantitative Analysis. 48, (2013), pp. 1001-1024.

Poledna, S., Molina-Borboa, J. L., Martínez-Jaramillo, S., van der Leij, M., Thurner, S. “The multi-layer network nature of systemic risk and its implications for the costs of financial crises.” Journal of Financial Stability. 20, (2015), pp. 70-81.

Jarrow, R. A., Lando, D., Turnbull, S.M. “A Markov model for the term structure of credit risk spreads.” Review of Financial Studies. 10, (1997), pp. 481-523.

Kleijnen, J. P. “Supply chain simulation tools and techniques: a survey.” International Journal of Simulation and Process Modelling. 1, (2005), pp. 82-89.

Kumiega, A., Neururer, T., Van Vliet, B. “Trading system capability.” Quantitative Finance. 14, (2014), pp. 383-392.

Leemis, L.M. Reliability: Probabilistic Models and Statistical Methods. (1995). Prentice Hall, Englewood Cliffs, NJ.

Markets Media. "NYSE outage aftermath barely registers." (2015). Accessed on July 30, 2015 at: marketsmedia.com/nyse-outage-aftermath-barely-registers/?utm_source=rss&utm_medium=rss?email=feat

Schuler, H. "Automation in chemical industry (Automatisierung in der chemischen industrie)." at-Automatisierungstechnik. 54, (2006), pp. 363-371.

Tomlin, B. "On the value of mitigation and contingency strategies for managing supply chain disruption risks." Management Science. 52, (2006), pp. 639-657.

Van Vliet, B., Cooper, R., Kumiega, A., Northey, J. "The rationale for AT 9000: An ISO 9000-style quality management system standard for automated and algorithmic trading." Journal of Trading. 8, (2013), 102-106.

Wall Street Journal (WSJ). "Loss swamps trading firm." (2012). Accessed on July 30, 2015 at: www.wsj.com/articles/SB10000872396390443866404577564772083961412

Wang, W., Loman, J., Vassiliou, P. "Reliability importance of components in a complex system." Reliability and Maintainability, 2004 Annual IEEE Symposium-RAMS. (2004), pp. 6-11.