Bank i Kredyt 49(1), 2018, 1-16

BANK I KREDYT
BANK & CREDIT

Toxic liquidity – is it here to stay?

Carlos Jorge Lenczewski Martins*

Submitted: 14 November 2017. Accepted: 3 January 2018.

Abstract

Algorithmic trading and high-frequency trading (HFT) volume in the financial market has dynamically increased mainly since the technological revolution of the 2000's. Turnover from HFT alone may be around 40% and 60% for the European and US equity markets, accordingly. The activity of algorithmic and high-frequency trading has also caused some concerns related to toxic liquidity in the financial market. These concerns reached their peak after the Flash Crash of 6 May 2010 in the US. After some reports of the Flash Crash reached the conclusion that what triggered the freefall was an algorithm, there were questions on whether HFT traders could induce toxic liquidity.

The aim of this paper is to present sources and conditions for which toxic liquidity may arise, as well as arguments that high-frequency traders may not necessarily be the source of toxic liquidity. Not only can they be a legitimate source of natural liquidity but also toxic liquidity may actually be attributable to low frequency traders. This paper is based on scarce empirical research available (mainly for the US market) and as such, it is intended to encourage further analysis and research on this important topic.

Keywords: algorithmic trading, high-frequency trading, HFT, liquidity

JEL: G11, G14, G15

* Warsaw School of Economics, Institute of Risk and Financial Markets; e-mail: clenc@sgh.waw.pl.

1 Introduction

The presence of HFT in the financial market is somewhat controversial. Some papers put HFT in the spotlight by stating that this type of trading negatively influences financial markets, while other papers state exactly the opposite – that positive changes in the markets are exactly derived from the presence of HFT. These discussions principally revolve around such market elements as volatility, spread or reaction time that market participants (low frequency traders, LFT) may not have in comparison to HFT firms.

Toxic liquidity is a broader concept that in fact takes into consideration all the above elements. Toxic liquidity may be shortly labelled as the situation when firms lead other firms to incur losses they otherwise wouldn't incur, by supplying them liquidity. Since many HFT firms act as market makers, there are also some doubts as to the whether they introduce toxic liquidity to the market or not. This suspicion was brought forward after one of the most dangerous events observed in the financial markets – the Flash Crash of May 2010 under which the presence of algorithmic trading was heavily active. During this event, the American markets plummeted by 10% in just 3 minutes, what of course, raised questions on how this happened and how to avoid the recurrence of such events in the future. Despite reports showing that HFT firms were not responsible for such incidents, some still find HFT responsible for losses incurred by many participants attributing them to toxic liquidity.

This paper is aimed first to present the characteristics of high-frequency trading and along the way to present some remarks on how HFTs are defined as well. The reason for this step is that there is still a misconception about what HFT is, and what it is not. The next goal is to differentiate volume, volatility and liquidity. This seems important for many reasons. First, because they do not necessarily follow from one another, i.e. volume does not necessarily lead to volatility and volatility is not a measure of liquidity. This also seems to be a concept that is sometimes mistaken. The second reason is related to, of course, toxic liquidity and the Flash Crash event which will be described in some detail simply because some papers relate the Flash Crash event to toxic liquidity as such. The final goal of this paper is also to discuss if HFT have influence over the toxic liquidity process, the potential spill-over effect to other market segments (e.g. the futures markets). It is worth mentioning that this paper is based on a limited number of papers available, and as such it is intended to encourage further research on this topic.

2 High-frequency trading and its effects on the financial market

The concept of high-frequency trading is directly connected with algorithmic trading which in turn is a type of electronic trading. While HFT can be considered to have been born in the 1990's, electronic trading is somewhat older. As such, HFT is the result of several steps taken towards the development of electronic trading. The beginnings of electronic trading go back to 1971 and 1976 (Agarwal 2012). In 1971, the National Association of Securities Dealers (NASD) exchange, which from 2001 functions as the National Association of Securities Dealers Automated Quotations (NASDAQ), introduced the possibility of electronic price quotations. It was not until 1976 that electronic trading entered a revolutionary expansion. That year, NYSE introduced in turn the possibility of buying and selling equities electronically. Another extremely important step was when the Securities and Exchange

Commission (SEC) in 1998 allowed the introduction of electronic exchanges in order to compete with traditional exchanges (Duhigg 2009). These steps were crucial not only to electronic trading and thus HFT, but also led to important structural and systemic changes of the financial market as well. Not only new entities emerged, but also new trading mechanisms between them. Although, the last decade of the 20th century may be considered the beginning of HFT, it was not until the technological revolution of the 2000's that HFT expanded dynamically. Research shows that in 2005 in the US the share of HFT in total equity trading already stood at 20%, while in Europe it was virtually non-existent (Kaya 2016). In just four years the share of HFT in the US was already 60%, against 40% for Europe in 2010 (Kaya 2016). Currently, the HFT share in the US is somewhat around 50% even after the decline resulting from the latest crisis (Goldstein, Kumar, Graves 2014).

In order to properly define HFT it seems reasonable to first define algorithmic trading (AT) since the first derives from the latter. There are many different definitions of AT. One concise definition runs as follows: "[…] the use of computer algorithms to automatically make trading decisions, submit orders, and manage those orders after submission." (Hendershott, Riordan 2009, p. 2). Other definitions, although similar, may include expressions that may not necessarily be accurate, such as the use of "pre-defined trading decisions" (Jarnecic, Snape 2010) in Gomber et al. (2011). Algorithmic trading does not necessarily mean that trading decisions are pre-defined. In fact, trading decisions are variable and depend on the result of information processing. It will all depend on the level of automation, or the complexity level of the algorithms used. The most advanced algorithms may select financial instruments, amounts, and the adequate moment for order submission independently with respect to opening or closing positions, changing orders or cancelling them. In fact, the Bank for International Settlements differentiates the level of automation used in the trading process (BIS 2011) contrasting algorithmic execution with algorithmic trade decision-making. In the first case, a human trader decides what to trade but then uses an algorithm to execute the trade. In the second type of AT, the trading process is completely automated, i.e. all the trading steps are automated. What this means is that the (algorithmic) model processes information, takes adequate trading decisions, and decides on how to execute the trade. There may be some predefined parameters but not decisions. These models may focus on order book imbalance, momentum, news trading, etc. but the actual decisions (as the name implies) are made by algorithms. This is an important distinction because it allows to categorise various types of AT, including high-frequency trading which falls under the scope of the second type of AT. Although segregating HFT into the categories of AT is an easy task, it seems that defining High-Frequency Trading is not. This is because many of the definitions rely only on some aspects of HFT and not on the "global" essence of HFT. One example may be "[…] the use of high--speed computer algorithms to automatically generate and execute trading decisions for the specific purpose of making returns on proprietary capital." (Jarnecic, Snape 2010) in Gomber et al. (2011, p. 76). In the paper by Jovanovic and Menkveld, HFT definition states that "computer algorithms are used […] to simply profit from buying and selling securities […]". While this is not an incorrect definition, it does not differentiate HFT from other traders. Other definitions rely on equating HFT with speed and collocation, which is generally correct, but interestingly enough not all HFT strategies imply special speed requirements and collocation (Narang 2010). Taking the approach by Gomber et al. (2011), it may be more reasonable to present the characteristics implied by HFT instead of defining it, since there seems to be no adequate definition for HFT. As for the entities concerned, HFT involves proprietary trading, with HFTs acting as middlemen – therefore profiting from buying and selling (market making). As for trading, it involves an outstandingly high number of orders, very short holding periods or rapid/frequent order cancellation. As to instruments, HFT focuses on highly liquid instruments, and since the holding periods are extremely short, margins per trade are consequently small. HFT firms in general have also a low latency requirement what means that they will use the latest technology and collocation or special data feeds. One last typical characteristic of HFT that bears mentioning is that they will be flat on positions by the end of the day.

In terms of regulatory definitions, it is the Directive 2014/65/EU on markets in financial instruments (MiFID II) that includes a definition of HFT for the first time. To a large extent it is in line with the previously presented definitions. Comment (61) of the MiFID, defines HFT as "[…] a trading system [that] analyses data or signals from the market at high speed and then sends or updates large numbers of orders within a very short time period in response to that analysis. In particular, high-frequency algorithmic trading may contain elements such as order initiation, generating, routing and execution which are determined by the system without human intervention for each individual trade or order, short time-frame for establishing and liquidating positions, high daily portfolio turnover, high order-to-trade ratio intraday and ending the trading day at or close to a flat position. High-frequency algorithmic trading is characterised, among others, by high message intra-day rates which constitute orders, quotes or cancellations". Additionally, "[HFT] rather than being a strategy in itself is usually the use of sophisticated technology to implement more traditional trading strategies such as market making or arbitrage" (Directive 2014/65/EU). Although, the definition used seems adequate, two significant remarks should be made related to "high message intra-day rates" and to the last statement that HFT is basically the use of sophisticated technology to implement more traditional trading strategies. The first issue is the problem of delineating the high message intra-day rates. The second issue is especially important because, although HFT is not a (single) strategy by itself and uses also "typical" strategies – it must not be regarded as low-frequency trading "on steroids" (Easley, Lopez de Prado, O'Hara 2012b) or "cheetah traders" (Lenzner 2011). There are many different business models (sometimes described as strategies) used by HFT including the "traditional" strategies such as arbitrage, order-book imbalances, market making, and others (Chlistalla 2011). The reality is, as usual, a bit different. What for LFT's is considered strategy, for HFT's should be considered a business model. The strategy(ies) behind these business models is another element that differentiates HFT's from LFT's. In addition, HFT firms very often use predatory trading strategies in order to detect other algorithms (HFT's or not) and limit orders placed by both LFT and HFT traders. These strategies may be very aggressive depending on different factors but HFT traders should be considered as extremely strategic traders (De Prado 2012).

As far as the influence of High-Frequency Trading in the financial market is concerned, research papers focus on such elements as volatility, Bid-Ask spreads, price discovery or adverse selection, but they are not particularly consistent in their findings. For example, Furse et al. (2011, p. 12) show "[…] no direct evidence that high frequency computer based trading has increased volatility". Also, Brogaard, Hendershott and Riordan (2014, p. 2303) find that there is "[…] no direct evidence that HFT contribute to market instability in prices". Another study by Chaboud et al. (2009, Abstract) in the exchange market also shows that "[…] there is no evident causal relationship between algorithmic trading and increased exchange rate volatility. If anything, the presence of more algorithmic trading is associated with lower volatility". Going even further, other studies show that the presence of HFT's improves market quality by increasing liquidity, lowering adverse selection, lowering costs of trading (e.g. the spread) and by increasing the informativeness of quotes and prices (Hendershott, Jones, Menkveld 2011), with respect to algorithmic trading or Malinova, Park and Riordan (2013). On the other hand, considering the same factors, other papers find opposite results. For example, Biais, Foucault and Moinas (2011) find that HFT generate adverse selection simply because they can process information before LFT. Boehmer, Fong and Wu (2015) find that AT "improves liquidity and informational efficiency but increases volatility" and "that these volatility-seeking traders are associated with declines in market quality".[1] Foucault, Hombert and Roşu (2016) state that when the trader is fast (HFT), the market may be less liquid, which increases adverse selection due to faster speeds or more precise information obtained (before others). Since technology in HFT is evolving, leading to greater possible speeds, when considering higher speeds than milliseconds or microseconds the results of Gai, Yao and Ye (2013, p. 1) show that the increase in "[…] the speed of trading from microseconds to nanoseconds do not lead to improvements on quoted spread, effective spread, trading volume or variance ratio" and "short term volatility increases and market depth decreases".

A report on HFT and AT from the Swedish Supervisory Authority from 2012, has stated that the "[…] impact of high frequency trading on financial stability is still limited" and "[…] the harmful effects […] are less extensive that many feared" (Finansinpektionen 2012, p. 4), but "[…] there is still considerable concern about market abuse" (Finansinpektionen 2012, p. 3), which brings us to another topic – market abuse and manipulation. An often voiced accusation and concern is that the presence of HFT leads to increase in market manipulation as a result of high-speed strategies that other market participants may be unaware of. The full list of market manipulation techniques (not only under HFT) is not the scope of this paper but it is important to focus on this matter for a while. The most frequent types of such strategies that may distort markets are probably quote stuffing, spoofing, momentum ignition, or layering. Quote stuffing, as the name implies, through significant amount of order submissions and cancellations is aimed to generate a large number of new best bid and best ask prices. The reasons behind this technique may be to engage other participants (or algos) in submitting orders and to provide price congestions to slow transaction processing times of HFTs and SFTs (Tse, Lin, Vincent 2012a). Spoofing is in some ways similar to quote stuffing in that it takes advantage from speed. The concept behind spoofing is to give a wrong impression of prices to other participants. A "spoofer" will open a limit order far from the best Bid or best Ask (either before or after) submitting large amounts of opposite limit orders in order to give the impression that prices are moving accordingly. When prices start falling/rising the "spoofer" will cancel these orders and wait until the first orders are executed. A simple example is when a "spoofer" wants to buy an asset. He will first submit (significant) sell limit orders above the Ask price, giving the impression of a potential price descent. Due to the speed advantage, it is not a problem for the trader to cancel these orders if prices start rising instead of falling. Either way, in the end the trader will cancel these orders and will have executed a buy limit order once prices have fallen. Momentum ignition, in turn, is somewhat similar to spoofing and is aimed to create short-term mispricing and to achieve a "long--term" profit from that action (Tse, Lin, Vincent 2012b). Momentum ignition is not a short-term "effect" as spoofing, but rather instigated by short-term actions. The main idea behind momentum ignition is to lead other participants to enter the market, in directional transactions, in response to short term actions produced (e.g. quote stuffing or spoofing), and profiting from pre-positions. In other words,

[1] It is important to underline that AT is quite different from its subset HFT, taking significantly longer positions and using different strategies, which will be highlighted later.

the instigator will open a pre-position, use high-frequency order submission and cancelation, which leads to directional bets from other participants, and when price has changed enough, the instigator will close its pre-position. Layering can be considered a form of spoofing (Gomber et al. 2011), and relies on submitting the same type of orders (e.g. buy limit orders) but at different price levels and cancelled when liquidity is demanded by another trader (Breckenfelder 2013). In that way, layering gives the impression of false liquidity (interest) of an asset. A trader may for instance submit orders on one side of the order book and hidden orders[2] on the other side of the book. Interestingly enough, as for the influence of HFT in market manipulation, a paper by Cumming, Zhan and Aitken (2012) shows that, in fact, HFT has diminished the frequency of different types of end-of-day manipulation. Although the level of potential manipulation by HFT firms is not known, there are some market events specifically attributed to HFT activity, and these are flash crashes.

3 Flash Crash of 6 May 2010

Flash crashes are "simply" extremely short-termed market disruptions, usually associated with algorithmic trading that can have long-term consequences. They appear in the form of extremely sharp and fast price depreciations, following which prices pull back repeatedly to somewhat lower levels than before the drop, i.e. the sharp drop may be of a magnitude of 5% but then the prices rise and stay at the level of 1% below the original price. The time frame for the price drop is not defined but most often it can go from milliseconds to a couple of minutes. Johnson et al. (2012) call these events "ultrafast black swan events".[3] The authors in their study differentiate the time frame for crashes and for spikes as less than 650ms for the first, and less than 950ms for the second. Additionally, to qualify them as "black swan crashes" stock prices should tick down at least ten times and prices have to fall at least 0.8% (Johnson et al. 2012). Although the authors focused on flash crashes that had a duration of less than 1800ms, the price drop of probably the most discussed flash crash (the one of 6 May 2010) had a total duration of close to 3 minutes (with many "sub-crashes" involved), while the price recovery took close to 15 minutes. It is difficult to even compare the consequences of, for example, the latest financial crisis, which started in 2008, to those of a Flash Crash. However, what is a cause for concern is the short period in which such significant losses are incurred, which under a flash crash is a matter of minutes.

The Flash Crash of 6 May 2010 is probably one of the most mentioned arguments against HFT and AT. That day, as presented in Figure 1, the American Dow Jones saw the largest one-day drop of close to 1000 points, while E-mini S&P 500 saw a drop of 5%. Some of the arguments considered as being the causes of the Flash Crash of 6 May 2010 will be presented below.

The whole day of 6 May 2010 saw liquidity diminish while volatility and negative market sentiment increased fast mainly due to concerns about the European debt crisis (U.S. Commodity Futures Trading Commission 2010). Not only was there a selling pressure on some stocks but there were also riots in Athens, which increased the selling pressure even further. At 14:32 a mutual fund started an algorithm ("the" algorithm) to sell 75,000 E-Mini contracts. Valued at close to USD 4.1 billion, instead

[2] Catherine D'Hondt provides two very interesting papers on the process of using and detecting hidden orders, and how they may affect market depth and transparency (D'Hondt, De Winne, Francois-Heude 2004; De Winne, D'Hondt 2007).

[3] A valid discussion rises of the differentiation of black swan and dragon king events (those that may be quantified, and have some kind of predictability) (Sornette 2009).

of selling at manually defined moments, these contracts were sold at a predefined level of 9% of the trading volume visible (SEC 2010). While volume based algorithms are very often used,[4] the one from 6 May did not take into consideration either the price or the time of sale – one may understand that it was meant to sell 9% of the market volume at whatever time and at whatever price. The consequences of disregarding prices or time in this algorithm were, of course, significant because depending on volume, the algorithm is executed more quickly or more slowly. On 6 May it took only 20 minutes to sell close to USD 4.1 billion worth of contracts. As some papers show (SEC 2010), while the AT initiated the selling, HFT traders were the ones initially providing liquidity to the market by accumulating initially long positions (close to 3,300 contracts). One may see from Figure 1 that from around 14:43 the selling pressure increased visibly, when HFT firms started liquidating their short-term long positions. As such, they were competing with "the" algorithm for liquidity which would be provided by other market participants (non-HFT as well). The Securities and Exchange Report for this event shows that HFTs traded nearly 140,000 E-Mini contracts or over 33% of the total trading volume – a behaviour which is common for HFT (SEC 2010). Inside Figure 1, the small chart to the left also shows the buying and selling volume of E-Mini and SPY.[5] Although there was some buying volume, it was not enough to balance out the selling volume, and to avoid the sudden drop in prices minutes later. Additionally, there was a "hot potato" effect in play among HFTs between 14:41 and the actual bottom of the E-Mini prices. During these 4 minutes, there was practically no trader willing to provide liquidity – something that changed after the circuit breaker initiated by CME was triggered and prevented further price declines. Interestingly, "the" algorithm continued to sell contracts after the circuit breaker,[6] but this time there were already participants in the market buying what was being sold – a reason why prices might have started to rebound slowly. Some firms even reduced or halted trading completely when prices started rebounding (Hasbrouck 2015). The report from the SEC also signals processing issues from market makers which in normal market conditions would have continued to provide liquidity. These issues were related with the extremely large amount of orders submitted by other participants and fast moving prices. Market makers (with human intervention) were not able to assess all the information arriving before price changes, what also led to market makers submitting stub quotes,[7] complying with their obligations to provide quotations for both bid and ask prices (SEC 2010). This is an important finding because by using stub--quotes market makers are offering quotes at very distant levels, which under such conditions as seen on 6 May and with the use of circuit brokers could be reached by market or limit orders.

In conclusion, there was a number of factors that led to the Flash Crash of 6 May 2010 but it seems far-going to attribute the blame to HFTs merely because they were active participants (just as LFTs) and took hedging related decisions since the decision to close their positions due to unfavourable market conditions should not be treated as a deliberate act due to faster information processing leading other participants to losses (i.e. toxic liquidity).

---

[4] Volume based algorithms are very often known as volume-weighted average price (VWAP) which is designed to achieve an average price from a specified period. Since it is based on volume, price and time it will therefore avoid to execute significant orders which may affect prices.

[5] SPY is an index aimed to mirror the S&P 500 index.

[6] Circuit breakers are exchange mechanisms that stop trading of one or more equities to prevent extremely fast price drops.

[7] Stub quotes is a practice market makers use to continue to provide two-sided quotes but at extremely high or low levels, which they believe will not be reached.

4 Volume, volatility, liquidity and toxic liquidity

One important factor involving flash crashes and toxic liquidity effects is the relation between liquidity, volume and volatility. Although, this relationship is not new, it is important to underline it, since it is one of the fields of study in microstructure theories and in the analysis of HFT impact on financial markets. There are some key principles originating, e.g. in Black (1971) and Kyle (1985) related to the definition of liquidity. A market is considered liquid when three conditions are met (Foucault, Kadan, Kandel 2005): it is tight – i.e. spreads are small, deep – the price impact is small, and resilient – prices recover quickly after demand shocks. Spreads increase and decrease for different reasons, such as information asymmetry, inventory costs, or other. This may also be understood as the cost of turning a position over a short period of time. Depth is related to the size of limit orders available in the order book at a given price level. The more limit orders on diverse price levels, the more market breadth there is (Aldridge 2009). Lastly, resiliency may be somewhat dependent on the previous two conditions and may be understood as the speed with which prices recover from a random uninformative shock (Kyle 1985). One other way to understand resiliency, which becomes fundamental in AT and HFT strategies, is "[…] the ability to trade at a minimal price impact" (Eren, Ozsoylev 2006, p. 24).

When analysing these preconditions of a liquid market, one may understand that at least two relations are essential to remember, even though sometimes they are forgotten. First, in a market situation where high volume and substantial liquidity are present, price changes should be small. Secondly, when there is high volume but low liquidity, price changes will become considerable. As such, high volume is not a synonym of sudden and large price changes (volatility). When taking into account the evolution of the Flash Crash of 6 May 2010, even with the involvement of HFTs which (could) generate high volume a sudden price drop would not have occurred if liquidity had not been withdrawn.

There is one last conclusion, or hypothesis that can be drawn from the two above mentioned relations: the precondition for a flash crash could be that there is high volume but the only liquidity available comes from HFT participants. Since high volume and low liquidity increases volatility, there is a chance that when this liquidity is being withdrawn suddenly (by HFTs), prices will suddenly change which conforms with the "definition" of a flash crash. Although other flash crash situations are not under the scope of this paper, in the case of the Flash Crash of 6 May 2010, this condition seems not to be met. It is a fact that HFT participants were active (generating between 33% and 49% of the total trading volume), and they were initially acting as market makers, but they were not the only liquidity providers at the time. They started to draw back earlier than other participants. Between 14:41 and 14:44 (before the most sudden drop in prices) HFTs started to withdraw from their initial liquidity provision. After that, they generated only 49% of volume between 14:45:13 and 14:45:27 (SEC 2010). When looking at Figure 1, this is the period when the most sudden price drop occurred. In short, HFTs were a big part of the total trading volume, they started to draw back from liquidity provision before other participants but they were not the only ones to provide liquidity.

High-frequency trading gives enormous advantages, the most important being the speed of trading and the speed of information processing. These advantages bring questions on stability issues, due to the fact that HFT traders make trading decisions faster than other market participants. This becomes especially important when HFT traders also take the role of liquidity providers, but not only. Due to speed advantages, low-frequency traders may not be willing to provide liquidity due to the obvious risk of adverse selection. When orders are being directed by better informed participants to market makers, and as a result the later suffer losses, then one may refer to this practice as toxic liquidity or order flow toxicity (Easley, Lopez de Prado, O'Hara 2011). Consequently, due to "worse" information processing market makers may be unaware they are providing liquidity at a loss (Easley, Lopez de Prado, O'Hara 2012a). Attributing HFT participants with the initiative of leading market makers to toxic liquidity would seem logical, but not it is not necessarily true – as seen during the Flash Crash of 6 May 2010. As previously mentioned, HFT participants were in fact providing liquidity (therefore not leading to toxic liquidity), and when they closed their positions, they closed the previously opened positions related to liquidity provision. This means that they were not closing positions due to better information processing but due to risk related issues.

There is one more important observation to mention. A study regarding NASDAQ presented by Carrion (2013) shows that HFTs provide liquidity when spreads are wider and take liquidity when spreads are tighter, which suggests that HFTs provide liquidity when it runs short, and take liquidity when it is vastly abundant. This may be an important observation because if LFT were taking liquidity from market makers, and these feared (toxic) adverse selection, then spreads should be wider, which under the study of Carrion does not happen, i.e. HFTs take liquidity when spreads are tight. On the one hand, one could say that in this situation the risk of toxic liquidity is low because information asymmetry may also be low. Nevertheless, it is also difficult to say whether the reason for spreads being tighter is lower information asymmetry or lower inventory costs. On the other hand, studies such as the one by Brogaard, Hendershott and Riordan (2014) show that when HFTs are demanding liquidity, they bring information into the market, therefore impose adverse selection on a larger scale than LFTs. As such, they could be taking advantage of narrow spreads and low risk premia from liquidity providers to open directional positions which can lead unaware market makers into losses. It is also important to mention the fact that even if HFTs really increase adverse selection costs, they also disseminate information more quickly than LFTs do. This means that although adverse selection may be higher at the beginning, it will dissipate rather quickly, which should lead to lower adverse selection costs in the long-run. Even if HFTs lead to toxic liquidity this effect should be short termed, in contrast to the potential toxic liquidity introduced by LFTs from which information is obtained more slowly. Kirilenko and Lo (2013) remark that as far as the Flash Crash of 6 May 2010 is concerned, HFT participants were one of the principal types of participants that lead to price crashes. They argue that after reaching their critical inventory levels, they began selling quite aggressively when liquidity was rather scarce. But the problem that seems to be the most important for events such as the Flash Crash, is the use of volume-based algorithms (rather than HFTs). Even if at the time, no HFTs had been involved, the moment when participants start selling (for whatever reason), volume-based algorithms will increase their selling volume and maybe speed as. This leads other participants, LFTs included, to similarly increase the selling pressure. In the end, the selling pressure will escalate at a geometrical pace. HFTs have different business models, but the analysis of the Flash Crash, did not attribute volume-based models directly to HFTs. HFTs rather took mere inventory decisions when critical levels were reached.

Interestingly, the study by Van Ness, Van Ness and Yildiz (2017) show, in general, a negative relation between HFT and order flow toxicity. In addition, the relation between participants during toxic liquidity events is not that obvious. Using the VPIN methodology as presented by Easley, Lopez de Prado and O'Hara (2012a), Van Ness, Van Ness and Yildiz (2017) find that first, HFT trades with

HFTs and non-HFTs are not associated with order flow toxicity. Second, trades between non-HFTs are the main sources of order flow toxicity. The basic explanation behind these observations refers to information asymmetry. If there is a high level of information asymmetry between liquidity providers and other participants, order flow toxicity will increase and so will losses for liquidity providers. Since many studies show that HFTs decrease, in fact, information asymmetry and increase price discovery, this means that in the short-term information asymmetry and toxic liquidity will decrease, as will losses for liquidity providers. Of course, there is still the long-lived information asymmetry and long--term positions in which HFTs are not active but LFTs are. This could be the main reason for why toxic liquidity is generated by LFTs and not HFTs and why it may be observable over longer periods.

5 Spill-over effects of toxic liquidity

High-frequency trading brings also new challenges for the interconnection analysis of various markets and instruments. This is because HFT uses diverse strategies where various instruments are considered. These could include, for example, a price arbitrage between stocks and futures following almost immediately the order book imbalances that make price change in the futures markets – something that happened in the Flash Crash of 2010. Another more controversial technique includes the venue fade consisting in submitting orders in numerous exchanges, and as soon as one of them is executed, cancelling all other orders across venues. This raises issues regarding the so-called phantom liquidity or the liquidity that is not supposed to be actually provided. The problem related to phantom liquidity, apart from lack of confidence, is that investors and/or other algorithms may submit and execute orders as a consequence of the order book analysis available at the time. Haldane (2011) summarises at least three channels of contagion: between stock prices and derivatives, between different exchanges and trading platforms, and finally across stocks. The last channel results from analysis showing that algorithms "[…] tend to amplify cross-stock correlation in the face of a rise in volatility due to their greater use of algorithmic trend-following and arbitrage strategies" (Haldane 2011, p. 11).

It is worth mentioning that although these channels may be activated as price movement mechanisms, it does not mean that toxic liquidity will get spilled over by HFT entities. As mentioned previously, the study by Van Ness, Van Ness and Yildiz (2017) shows that HFT is not the main source of toxic liquidity while LFTs are. Taking this into account, one may see that a potential spill-over of toxic liquidity is possible, but not at a maximum potential speed, since its propagation would rather be induced by LFTs.

There is one element of contagion that is not directly related to toxic liquidity but could act as a trigger. This is the possibility that HFTs may induce contagion among traders, which in turn may lead other entities to introduce toxic liquidity. The literature listing HFT traders as one of the causes for flash crashes, also shows that these may induce contagion "[…] due to the correlated nature of securities and accelerated time scales and [due to the fact] that such mini crashes continuously occur in single stocks" (Lange, Borch 2014, p. 3). There is however, also an important aspect of "contagion" that is sometimes not mentioned, i.e. intention. Assuming for a moment that there exists something like contagion, the question is if this phenomenon is due to purposeful action or not. More specifically, the inquiry sets about not to ascertain if an entity decides to intentionally drive prices down (which could better be called manipulation), but rather if it engages in rational decision-making which affects other instruments, for example. An interesting approach to the notion of "contagion" by Easley, Lopez de Prado and O'Hara (2014) mentioned by Lange and Borch (2014, p. 7), treats it as a result "[…] of market makers hedging their risk of adverse selection in one instrument by taking liquidity in another, with the hope that over time they will be able to unwind their position in both instruments at a profit". This specific situation probably should not be treated as an intentionally driven contagion since it is purely a hedging decision not directed at other participants as in the case of toxic liquidity.

Nevertheless, firstly, there is no actual evidence that the different types of contagion constitute a specific result of HFT actions. Secondly, even if these consequences stemmed from HFT actions it may be difficult to argue that they are linked to a more malign intent other than a pure and simple hedging activity. Finally, because of both these arguments and since HFTs are less likely to generate toxic liquidity, it is also less likely that HFTs will induce toxic liquidity contagion.[8]

6 Summary

The presence of new trading mechanisms, especially the more aggressive ones, leads to more speculative and aggressive markets. As a result, this may also lead to more "predatory" trading strategies or such strategies that lead other participants to incur losses – something known as toxic liquidity or order flow toxicity.

Although the rise of algorithmic trading can be traced back to the 1990's, the knowledge on the influence of high-frequency trading (a subset of algorithmic trading) is still limited. There is some fair amount of studies on the relation between liquidity, volatility and market stability but the results are not consistent. This is especially true when tackling issues related to toxic liquidity. Most of the papers seem to burden HFT with the possibility of inducing toxic liquidity or even inducing contagion on the basis of the Flash Crash of 6 May 2010. In reality, HFTs initially provided liquidity and competed in this domain with a volume-based algorithm. After prices reached levels critical for the HFTs, they liquidated their positions – which is nothing abnormal. For some authors, it seems as if this could be interpreted as toxic liquidity, which it is not, since the action of selling assets was not intentionally directed at other participants with the intention to make them incur losses. Apart from the consideration of this market event dating back to 2010, there is at least one more study which shows that LFTs are more inclined to induce toxic liquidity than HFTs. Addressing the question whether toxic liquidity is here to stay, the answer is probably affirmative – it is and at a level not smaller than it has always been (due to LFTs actions). As for potential contagion or spill-over of toxic liquidity from HFTs, it seems difficult though to show that since HFTs are not inclined to induce toxic liquidity in one instrument or market, they might be even less probable to induce it across instruments or markets. Yet, it does not seem very probable. It is nevertheless possible for a "normal" spill-over or contagion to occur, induced either by HFTs or LFTs as a consequence of regular hedging activities.

[8] More general information on the different aspects of contagion (by HFT) may be found for example in Easley et al. (2014), Lange, Borch (2014) or Sornette, Von der Becke (2011).

References

Agarwal A. (2012), *High frequency trading: evolution and the future*, Capgemini, https://www.capgemini.com/resource-file-access/resource/pdf/High_Frequency_Trading__Evolution_and_the_Future.pdf.

Aldridge I. (2009), *High-frequency trading: a practical guide to algorithmic strategies and trading systems*, John Wiley and Sons.

Biais B., Foucault T., Moinas S. (2011), *Equilibrium high frequency trading*, International Conference of the French Finance Association (AFFI), https://ssrn.com/abstract=1834344.

BIS (2011), *High-frequency trading in the foreign exchange market*, https://www.bis.org/publ/mktc05.pdf.

Black F. (1971), Toward a fully automated stock exchange. Part I, *Financial Analysts Journal*, 27(4), 28–35+44.

Boehmer E., Fong K.Y.L., Wu J.J. (2015), *International evidence on algorithmic trading,* AFA 2013 San Diego Meetings Paper, https://ssrn.com/abstract=2022034.

Breckenfelder J.H. (2013), *Competition between high-frequency traders, and market quality*, NYU Stern Microstructure Meeting, https://ssrn.com/abstract=2264858.

Brogaard J., Hendershott T., Riordan R. (2014), High-frequency trading and price discovery, *Review of Financial Studies*, 27(8), 2267–2306.

Carrion A. (2013), Very fast money: high-frequency trading on the NASDAQ, *Journal of Financial Markets,* 16(4), 680–711.

Chaboud A., Chiquoine B., Hjalmarsson E., Vega C. (2009), Rise of the machines: algorithmic trading in the foreign exchange market, *The Journal of Finance*, 69(5), 2045–2084.

Chlistalla M. (2011), *High-frequency trading – better than its reputation?*, Deutsche Bank, https://c.mql5.com/forextsd/forum/168/high-frequency_trading–better_than_its_reputation.pdf.

Cumming D., Zhan F., Aitken M. (2012), *High frequency trading and end-of-day manipulation*, https://ssrn.com/abstract=2109920.

D'Hondt C., De Winne R., Francois-Heude A. (2004), *Hidden orders on Euronext: Nothing is quite as it seems*, Working paper, FuCaM-Catholic University of Mons.

De Prado M.L. (2012), *Low-frequency traders in a high-frequency world: a survival guide, computational research division*, https://ssrn.com/abstract=2150876.

De Winne R., D'Hondt C. (2007), Hide-and-seek in the market: placing and detecting hidden orders, *Review of Finance*, 11(4), 663–692.

Duhigg C. (2009), Stock traders find speed pays, in milliseconds, *NY Times*, 23 July, http://www.nytimes.com/2009/07/24/business/24trading.html?_r=0.

Easley D., López de Prado M., O'Hara M. (2011), The microstructure of the "Flash Crash": flow toxicity, liquidity crashes and the probability of informed trading, *The Journal of Portfolio Management*, 37(2), 118–128.

Easley D., de Prado M.M.L., O'Hara M. (2012a), Flow toxicity and liquidity in a high-frequency world, *Review of Financial Studies*, 25(5), 1457–1493.

Easley D., López de Prado M., O'Hara M. (2012b), The volume clock: insights into the high-frequency paradigm, *The Journal of Portfolio Management*, 39(1), 19–29.

Easley D., López de Prado M., O'Hara M. (2014), *High-frequency Trading: New Realities for Traders, Markets and Regulators*, Risk Books.

Eren N., Ozsoylev H.N. (2006), *Hype and dump manipulation*, AFA 2008 New Orleans Meetings Paper, http://www.sabanciuniv.edu/HaberlerDuyurular/Documents/DD20070403110757/rumorNov30.pdf.

Finansinpektionen (2012), *Investigation into high frequency trading and algorithmic trading, FI Report,* February, http://www.fi.se/contentassets/a9e47d166ba4466586bad3f216b46355/htf_eng.pdf.

Foucault T., Hombert J., Roşu I. (2016), News trading and speed, *The Journal of Finance*, 71(1), 335–382.

Foucault T., Kadan O., Kandel E. (2005), Limit order book as a market for liquidity, *The Review of Financial Studies*, 18(4), 1171–1217.

Furse C. et al. (2012), *The future of computer trading in financial markets*, Working paper, The Government Office for Science, London http://www.cftc.gov/idc/groups/public/@aboutcftc/documents/file/tacfuturecomputertrading1012.pdf.

Gai J., Yao C., Ye M. (2012), *The externalities of high frequency trading*, https://www.sec.gov/divisions/riskfin/seminar/ye031513.pdf.

Goldstein M.A., Kumar P., Graves F.C. (2014), Computerized and high frequency trading, *Financial Review*, 49(2), 177–202.

Gomber P., Arndt B., Lutat M., Uhle T. (2011), *High-frequency trading*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1858626.

Haldane A. (2011), *The race to zero*, speech at International Economic Association Sixteenth World Congress, Beijing, http://www.bis.org/review/r110720a.pdf.

Hasbrouck J. (2015), *High-frequency trading (HFT) & flash crashes*, http://people.stern.nyu.edu/jhasbrou/Teaching/POST%202015%20Fall/classNotes/High-Frequency%20Trading%20(HFT).pdf.

Hendershott T., Jones C.M., Menkveld A.J. (2011), Does algorithmic trading improve liquidity, *The Journal of Finance*, 66(1), 1–33.

Hendershott T., Riordan R. (2009), *Algorithmic trading and information,* http://faculty.haas.berkeley.edu/hender/ATInformation.pdf.

Jarnecic E., Snape M. (2010), *An analysis of trades by high frequency participants on the London Stock Exchange*, 17th Annual Conference of the Multinational Finance Society, MFS.

Johnson N., Zhao G., Hunsader E., Meng J., Ravindar A., Carran S., Tivnan B. (2012), *Financial black swans driven by ultrafast machine ecology*, https://arxiv.org/pdf/1202.1448.pdf.

Kaya O. (2016), *High-frequency trading: reaching the limits*, Deutsche Bank, https://www.dbresearch.com/PROD/RPS_EN-PROD/PROD0000000000454703/Research_Briefing%3A_High-frequency_trading.pdf.

Kirilenko A.A., Lo A.W. (2013), Moore's law versus Murphy's law: algorithmic trading and its discontents, *Journal of Economic Perspectives*, 27(2), 51–72.

Kyle A.S. (1985), Continuous auctions and insider trading, *Econometrica*, 53(6), 1315–1335.

Lange A., Borch C. (2014), *Contagious markets: on crowd psychology and high-frequency trading*, http://www.nanex.net/aqck2/Contagious.Markets.pdf.

Lenzner R. (2011), "Cheetah" traders are dangerously controlling our markets, *Forbes*, 8 October, https://www.forbes.com/sites/robertlenzner/2011/10/08/cheetah-traders-are-dangerously-controlling-our-markets/#544c8f607094.

Malinova K., Park A., Riordan R. (2013), *Do retail traders suffer from high frequency traders*, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2183806.

Nagel J. (2012), *High-frequency trading and market implications – an assessment from a central bank perspective*, Deutsche Bundesbank, https://www.bundesbank.de/Redaktion/EN/Downloads/Press/Presentations/2012_07_04_nagel_hft_und_martkimplikationen.pdf?__blob=publicationFile.

Narang M. (2010), *Tradeworx, Inc. public commentary on SEC market structure concept release*, https://www.sec.gov/comments/s7-02-10/s70210-129.pdf.

SEC (2010), *Findings regarding the market events of May 6, 2010*, Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues, https://www.sec.gov/news/studies/2010/marketevents-report.pdf.

Sornette D. (2009), Dragon-kings, black swans and the prediction of crises, *International Journal of Terraspace Science and Engineering*, 2(1), 1–18.

Sornette D., Von der Becke S. (2011), *Crashes and high frequency trading: An evaluation of risks posed by high speed algorithmic trading*, Research Paper Series, 11(63), Swiss Finance Institute.

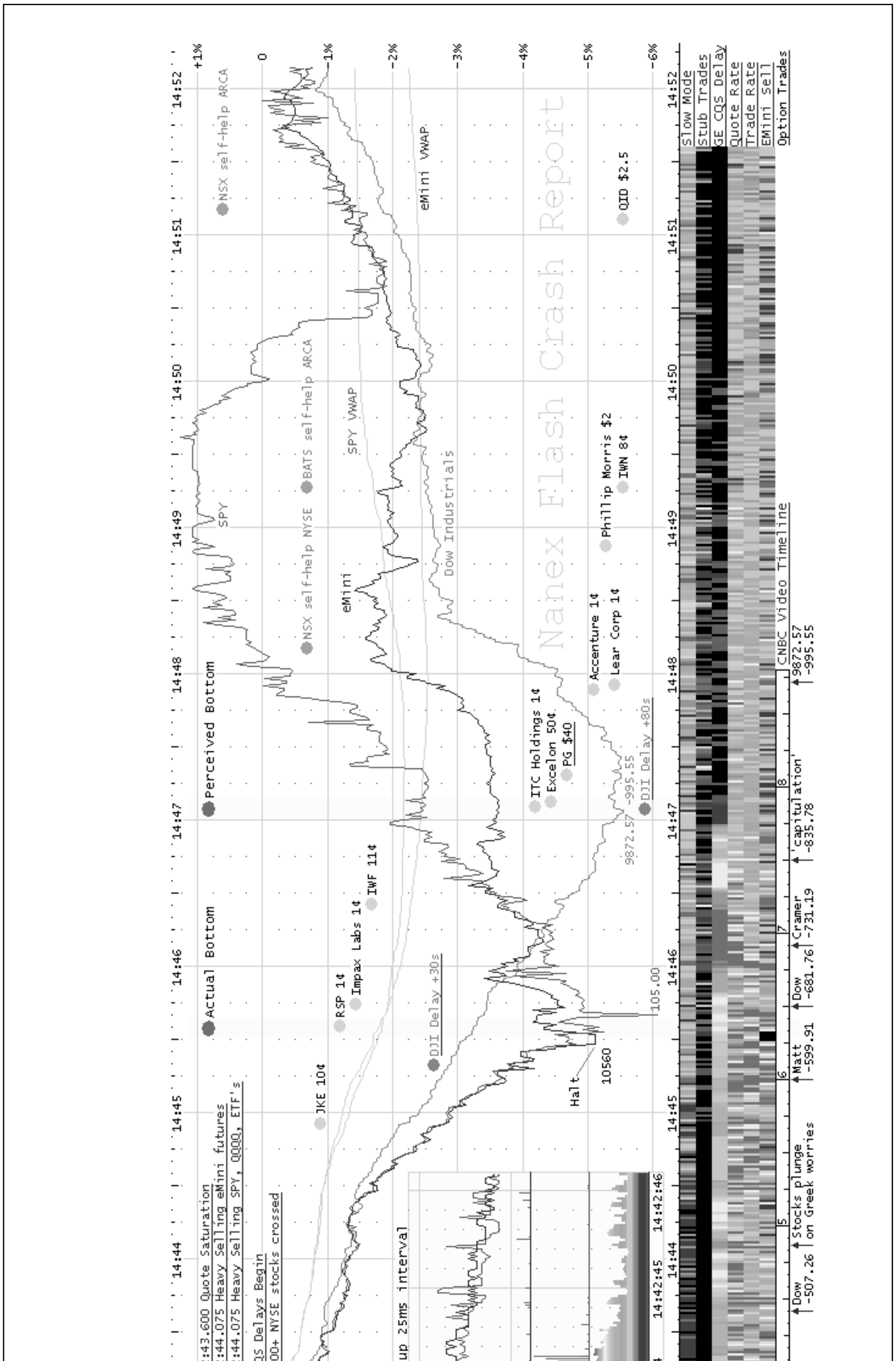
Tse J., Lin, X., Vincent D. (2012a), *High frequency trading – measurement, detection and response*, Credit Suisse, https://edge.credit-suisse.com/edge/Public/Bulletin/Servefile.aspx?FileID=23285&m=923226224.

Tse J., Lin X., Vincent D. (2012b), *High frequency trading – the good, the bad, and the regulation*, Credit Suisse, https://edge.credit-suisse.com/edge/Public/Bulletin/Servefile.aspx?FileID=23284&m=1815212669.

U.S. Commodity Futures Trading Commission (2010), *Preliminary 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, https://www.sec.gov/sec-cftc-prelimreport.pdf.

Van Ness B.F., Van Ness R.A., Yildiz S. (2017), The role of HFTs in order flow toxicity and stock price variance, and predicting changes in HFTs' liquidity provisions, *Journal of Economics and Finance*, 41(4), 739–762.

Figure 1
Quotations at one-second intervals from 14:42:30 to 14:52:00, of the Flash Crash of 6 May 2010 as seen in the US



Source: (Nagel 2012) from: Nanex.